

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу**

До захисту допущено:

Завідувач кафедри

_____ Оксана ТИМОЩУК

«___» _____ 20__ р.

Дипломна робота

на здобуття ступеня бакалавра

**за освітньо-професійною програмою «Інтелектуальний аналіз даних в
управлінні проектами» спеціальності 122 «Комп'ютерні науки та
інформаційні технології»
на тему: «Моделювання і прогнозування процесів у телекомунікаціях»**

Виконав (-ла):

Студент (-ка) IV курсу, групи КА-66

Абдал-Бакі Аміна Мутхуфарівна _____

Керівник:

професор, д.т.н., професор кафедри ММСА

Бідюк Петро Іванович _____

Консультант з економічної частини:

доцент, к.е.н. Шевчук Олена Анатоліївна _____

Консультант з нормоконтролю:

доцент, к.т.н. Коваленко Анатолій Спіфанович _____

Рецензент:

професор, д.т.н., професор кафедри ІБ ФТІ

Архипов Олександр Євгенович _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент (-ка) _____

Київ – 2020 рік

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 122 «Комп’ютерні науки та інформаційні технології»

Освітньо-професійна програма «Інтелектуальний аналіз даних в управлінні проектами»

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри

_____ Оксана ТИМОЩУК

«25» травня 2020 р.

ЗАВДАННЯ
на дипломну роботу студенту
Абдал-Бакі Аміні Мутхуфарівні

1. Тема роботи «Моделювання і прогнозування процесів в телекомунікаціях», керівник роботи Бідюк Петро Іванович, д.т.н., професор, затверджені наказом по університету від «25» травня 2020 р. № 1143-с
2. Термін подання студентом роботи 8.06.2020.
3. Вихідні дані до
4. Зміст роботи
5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)

6. Консультанти розділів роботи*

| Розділ | Прізвище, ініціали та посада консультанта | Підпис, дата | |
|-------------|---|----------------|------------------|
| | | завдання видав | завдання прийняв |
| Економічний | Шевчук О.А., доцент | | |

7. Дата видачі завдання _____

* Якщо визначені консультанти. Консультантом не може бути зазначено керівника дипломної роботи.

Календарний план

| № з/п | Назва етапів виконання дипломної роботи | Термін виконання етапів роботи | Примітка |
|-------|--|--------------------------------|----------|
| 1 | Формулювання тематики (напрямку) дослідження. | 03.09.2019 – 30.09.2019 | |
| 2 | Аналіз актуальності задач стосовно тематики дослідження | 01.10.2019 – 30.10.2019 | |
| 3 | Аналіз відомих результатів стосовно тематики дослідження | 01.11.2019 – 30.11.2019 | |
| 4 | Формулювання задач дослідження | 01.12.2019 – 30.12.2019 | |
| 5 | Уточнення теми дипломної роботи | 25.02.2019 | |
| 6 | Збір статичних даних, попередній аналіз даних | 01.03.2020 – 30.03.2020 | |
| 7 | Розробка програмного продукту для виконання обчислювальних експериментів | 01.03.2020 – 30.04.2020 | |
| 8 | Виконання обчислювальних експериментів, аналіз та оформлення результатів | 01.05.2020 – 20.05.2020 | |
| 9 | Оформлення пояснювальної записки у цілому | 21.05.2020 – 31.05.2020 | |
| 10 | Підготовка презентації для захисту | 28.05.2020 – 01.06.2020 | |
| 11 | Попередній захист дипломної роботи | 01.06.2020 – 03.06.2020 | |
| 12 | Захист дипломної роботи | 15.06.2020 – 19.06.2020 | |

Студент

Аміна АБДАЛ-БАКІ

Керівник

Петро БІДЮК

РЕФЕРАТ

Дана дипломна робота містить 99 ст., 4 ч., 8 табл., 36 рис., 3 дод., 11 джерел.

МОДЕЛЮВАННЯ, ПРОГНОЗУВАННЯ, ТЕЛЕКОМУНІКАЦІЇ,
АВТОРЕГРЕССИЯ, МОДЕЛЬ БОКСА-ДЖЕНКІНСА

Тема: Моделювання і прогнозування процесів у
телекомунікаціях.

У роботі розв'язується задача математичного моделювання і короткострокового прогнозування нестационарних процесів у телекомунікаціях.

Об'єкт дослідження: сучасна критеріальна база - множина статистичних критеріїв якості

Предмет дослідження: засоби моделювання і короткострокового прогнозування з застосуванням моделі Бокса-Дженкінса.

Мета роботи: дослідити ринок телекомунікацій в Україні та спрогнозувати майбутнє монополіста на цьому ринку – ПАТ «Укртелеком»

Методи дослідження: використано математичний апарат регресійного моделювання, статистичні дані стосовно розвитку вибраних процесів у телекомунікаційній сфері

Створено програмний продукт для прогнозування процесів в сегменті B2B та B2C в телекомунікаціях. Для проведення аналізу було використано реальні дані кількості підключення та відключення фіксованого зв'язку та інтернету ПАТ «Укртелеком» в період з січня 2019 до грудня 2019.

ABSTRACT

This thesis contains 99 pages, 4 sections, 8 tables, 36 figures, 3 appendices, 11 sources.

SIMULATION, FORECASTING, TELECOMMUNICATIONS, AUTOREGRESSION, ARIMA MODEL.

Theme: Modeling and forecasting of processes in telecommunications.

The problem of mathematical modeling and short-term forecasting of non-stationary processes in telecommunications is solved in the work..

Object of research: modern criterion base - a set of statistical quality criteria.

Subject of research: means of modeling and short-term forecasting using the ARIMA model.

The purpose of the work is to study the telecommunications market in Ukraine and predict the future of the monopolist in this market - PJSC "Ukrtelecom"

Methods of research: used mathematical apparatus of regression modeling, statistical data on the development of selected processes in the telecommunications sector

A software product for forecasting processes in the B2B and B2C segment in telecommunications has been created. For the analysis, real data on the number of connection and disconnection of fixed communication and Internet of PJSC "Ukrtelecom" in the period from January 2019 to December 2019 were used.

ЗМІСТ

| | |
|--|-----------|
| ВСТУП..... | 9 |
| РОЗДІЛ 1 АКТУАЛЬНІСТЬ ДОСЛІДЖЕННЯ, ІСНУЮЧІ СИСТЕМИ ДЛЯ МОДЕЛЮВАННЯ ПРОГНОЗІВ ФІНАНСОВО-ЕКОНОМІЧНИХ ПРОЦЕСІВ | 11 |
| 1.1 Теоретичні відомості телекомунікаційної сфери | 11 |
| 1.1.2 Еволюція телекомунікаційного сектора..... | 11 |
| 1.1.3 Як телекомунікаційні компанії заробляють гроші?..... | 13 |
| 1.2 Дослідження нинішньої ситуації на телекомунікаційному ринку у світі. Ключові сегменти телекомунікаційної галузі | 13 |
| 1.2.1 Інвестування в телекомунікації..... | 14 |
| 1.2.2 Оцінка міжнародних телекомунікаційних компаній | 15 |
| 1.2.3 Телекомунікаційні ETFs та огляд телекомунікаційного сектора..... | 18 |
| 1.3 Рейтинг українських інтернет-провайдерів | 20 |
| 1.4 Актуальність теми дослідження та визначення існуючих проблем.... | 21 |
| 1.5 Висновки до 1 розділу..... | 24 |
| РОЗДІЛ 2 МЕТОДИ МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ ПРОЦЕСІВ НА ТЕЛЕКОМУНІКАЦІЙНОМУ РИНКУ УКРАЇНИ | 25 |
| 2.1 Дослідження телекомунікаційного ринку України. Підходи до аналізу | 25 |
| 2.2 Аналітична частина | 28 |
| 2.3 Існуючі моделі прогнозування | 34 |
| 2.4 Програмне забезпечення для прогнозування, переваги та недоліки | 42 |
| Висновки до розділу 2 | 55 |
| РОЗДІЛ 3 РОЗРОБКА ПРОГРАМНОГО ПРОДУКТУ ТА ЙОГО ЗАСТОСУВАННЯ..... | 56 |
| 3.1 Розробка програмного продукту | 56 |
| 3.2 Вибір інструментальної платформи для реалізації програмного продукту, інструкція по експлуатації програми | 57 |
| 3.3 Інструкція по експлуатації..... | 58 |
| Висновки до розділу 3 | 70 |

| | |
|---|-----------|
| РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ | 73 |
| 4.1 Постановка задачі | 73 |
| 4.1.1 Обґрунтування функцій програмного продукту | 74 |
| 4.2Обґрунтування системи параметрів ПП | 76 |
| 4.3Аналіз рівня якості варіантів реалізації функцій..... | 80 |
| 4.4 Економічний аналіз варіантів розробки ПП..... | 81 |
| 4.5 Вибір кращого варіанта ПП техніко-економічного рівня..... | 84 |
| Висновки до 4 розділу | 84 |
| ВИСНОВКИ | 86 |
| СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ..... | 87 |
| ДОДАТОК А..... | 88 |
| ДОДАТОК Б..... | 94 |
| ДОДАТОК В..... | 97 |

ВСТУП

В останнє десятиліття спостерігається швидкий прогрес у телекомунікаційних технологіях на все більш дерегульованому та конкурентному ринку. Компанії, що працюють на різних ринках, покладалися на прогнози попиту, щоб обґрунтувати значні інвестиції, необхідні для забезпечення доступності потужностей у потрібний час. Ці нові ринки, як правило, складаються з нових споживачів, які вперше приймають товар чи послугу, встановлених користувачів, що змінюють схему використання, користувачів конкуруючих послуг, що переходять на альтернативні послуги, та тих, хто взагалі виходить з цього сегмента ринку.

Прогнозування - один з ключових моментів у прийнятті управлінських рішень у постіндустріальному суспільстві. Регулярне прогнозування процесів дозволяє не тільки приймати ефективні управлінські рішення, а й накопичувати досвід, що дозволяє підвищити точність та надійність прогнозів, удосконалити методи та алгоритми прогнозування.

Методи прогнозування, що лежать в основі управління, потрібні для визначення загальної перспективи та тенденцій розвитку прогнозованих процесів, щоб забезпечити баланс між короткотерміновими та довгостроковими програмами. Більше того, важливо встановити сукупність факторів, що впливають на розвиток певних процесів.

Змінність виробничих та економічних відносин, нестабільність, концептуальна непослідовність та незавершеність законодавчої бази, що регулює ці відносини, необхідність адекватних управлінських рішень у нових сферах людської діяльності тощо призводять до того, що часові ряди, що представляють багато виробничі, економічні та соціально-економічні процеси, мають коротку актуальну частину. Крім того, багато нових процесів

у виробничій та соціальній сфері представляються фізично короткими часовими рядами, оскільки вони раніше не підлягали статистичному обліку.

У другому розділі розглянуто аналіз методів моделювання і прогнозування процесів на телекомунікаційному ринку України..

У третьому розділі подано процес проектування програмного продукту для аналізу процесів у телекомунікаційній сфері, результати виконання обчислювальних експериментів стосовно побудови математичних моделей вибраних процесів та оцінювання короткострокових прогнозів.

У четвертому розділі описано результати функціонально-вартісного аналізу програмного продукту.

РОЗДІЛ 1 АКТУАЛЬНІСТЬ ДОСЛІДЖЕННЯ, ІСНУЮЧІ СИСТЕМИ ДЛЯ МОДЕЛЮВАННЯ ПРОГНОЗІВ ФІНАНСОВО-ЕКОНОМІЧНИХ ПРОЦЕСІВ

1.1 Теоретичні відомості телекомунікаційної сфери

Сектор телекомунікацій складається з компаній, які роблять можливим спілкування в глобальному масштабі, по телефону або через Інтернет, через ефір або кабелі, по дротах або по бездротовому зв'язку. Ці компанії створили інфраструктуру, яка дозволяє відправляти дані в вигляді слів, голосу, аудіо або відео в будь-яку точку світу. Найбільшими компаніями в цьому секторі є оператори телефонного (як дротового, так і бездротового) зв'язку, супутникові компанії, кабельні компанії і постачальники інтернет-послуг.

Не так давно телекомунікаційний сектор складався з клубу великих національних і регіональних операторів. З початку 2000-х років галузь була охоплена швидкою дерегуляцією і інноваціями. У багатьох країнах світу державні монополії в даний час приватизовані і стикаються з безліччю нових конкурентів. Традиційні ринки були перевернуті з ніг на голову, оскільки зростання мобільних послуг випереджає фіксовану лінію, і інтернет починає замінювати голос, як основний бізнес.

1.1.2 Еволюція телекомунікаційного сектора

Телекомунікаційна індустрія почалася в 1830-х роках з винаходу телеграфу, першого механічного пристрою зв'язку. Це скоротило швидкість обміну даними з декількох днів до декількох годин - так само, як сучасні

мобільні технології скоротили часовий інтервал відправки великих обсягів даних з годин до секунд. Промисловість розширювалася з кожним новим винаходом: телефон, радіо, телебачення, комп'ютер, мобільний пристрій. Ці технологічні досягнення змінили те, як люди живуть і займаються бізнесом.

Свого часу телекомунікації вимагали фізичних проводів, що з'єднують будинки і підприємства. У сучасному суспільстві технології стали мобільними. В даний час бездротові цифрові технології стають основною формою спілкування.

Структура сектора також змінилася з декількох великих гравців на більш децентралізовану систему зі зменшеним регулюванням і бар'єрами для входу. Великі державні корпорації виступають в якості постачальників послуг, в той час як невеликі компанії продають і обслуговують обладнання, таке як маршрутизатори, комутатори і інфраструктуру, яка забезпечує цей зв'язок.

1.1.3 Як телекомунікаційні компанії заробляють гроші?

Звичайні телефонні дзвінки продовжують залишатися найбільшим джерелом доходу в галузі, але завдяки досягненням в сфері мережевих технологій ситуація змінюється. Телеком все менше дбає про голос і все більше про відео, тексти і дані. Основна технологія широкосмугового зв'язку, цифрова абонентська лінія (DSL), відкрила нову еру. Найшвидше зростання відбувається від послуг, що надаються через мобільні мережі.

Успіх телекомунікаційних компаній багато в чому залежить від сили бренду і великих інвестицій в ефективні біллінгові системи. Корпоративний ринок, з іншого боку, залишається фаворитом галузі. Великі корпоративні клієнти, які в основному стурбовані якістю і надійністю своїх телефонних дзвінків і доставкою даних, менш чутливі до цін, ніж побутові клієнти. Наприклад, великі транснаціональні корпорації витрачають великі кошти на телекомунікаційну інфраструктуру для підтримки масштабних операцій. Вони також із задоволенням оплачують послуги преміум-класу, такі як приватні мережі з високим рівнем безпеки.

Оператори зв'язку також заробляють гроші, надаючи підключення до мережі іншим телекомунікаційним компаніям, які цього потребують, і продаючи канали для великих користувачів мережі, таких як інтернет-провайдерів та великих корпорацій. Взаємопов'язані і оптові ринки сприяють тим гравцям, які мають далекосяжні мережі.

1.2 Дослідження нинішньої ситуації на телекомунікаційному ринку у світі.

Ключові сегменти телекомунікаційної галузі

Сектор телекомунікацій складається з трьох основних підсекторів: телекомунікаційне обладнання (найбільше), телекомунікаційні послуги (наступне за величиною) і бездротовий зв'язок. Основні сегменти в цих підсекторах включають таке:

- Бездротовий зв'язок
- Устаткування зв'язку
- Внутрішні послуги зв'язку
- Зарубіжні телекомунікаційні послуги
- Диверсифіковані послуги зв'язку

Найменша, але найбільш швидкозростаюча область в цьому секторі - це бездротовий зв'язок, оскільки все більше і більше комунікаційних і обчислювальних методів переходять на мобільні пристрої і хмарні технології. Ця частина галузі є очікуваною для подальшого глобального розширення телекомунікаційного сектора. Є ще можливості для зростання навіть в розвинених країнах: наприклад, в 2018 році Федеральна комісія із зв'язку (FCC) повідомила, що приблизно одна п'ята частина сільського населення США як і раніше не має доступу до широкосмугових мереж.

1.2.1 Інвестування в телекомунікації

Телекомунікаційні компанії є рідкістю серед акцій: їх акції іноді демонструють характеристики як доходів, так і акцій зростання. Для зростаючих інвесторів невеликі компанії, що пропонують бездротові послуги, надають кращі можливості для підвищення ціни акцій. Навпаки, великі компанії, що займаються обладнанням і послугами, як правило, є «притулком» для консервативних, орієнтованих на дохід інвесторів.

Цінні інвестори також можуть знайти хороший вибір в телекомунікаційному секторі. Потреба в телекомунікаційних послугах, які є невід'ємною частиною світової економіки, зберігається незалежно від змін в діловому циклі. Однак, хоча попит постійний, окремі постачальники можуть рости і падати. Протягом декількох років компанія може користуватися своїми регулятивними привілеями (як і інші комунальні підприємства, телекомунікаційні компанії часто захищені від конкуренції за дорученням уряду) і отримувати надійні, щедрі дивіденди (генеруються високим щомісячним доходом від стабільної клієнтської бази). Потім, раптово, технологічний прогрес або злиття і поглинання створюють невизначеність і залишають місце для втрат - і відновлення з новим зростанням.

Якщо фірма стикається з різким спадом через зміни в галузі (наприклад, зростаючу важливість бездротових пристроїв), ціннісні інвестори можуть її поглинути, за умови, що її основи залишаються сильними і вона виявляється здатною адаптуватися до змін. Сектор телекомунікацій у виплаті та регулярне підвищення дивідендів робить період очікування поліпшення цін на акції більш приємним.

Проте, усі три основних телекомунікаційних сектора представляють певний ризик для інвесторів, оскільки акції реєструються в діапазоні від 7% (для послуг) до 15% (для бездротових мереж) і на 24% (для обладнання) з більшою волатильністю, ніж на більш широкому ринку. Інвестори, сильно піддані впливу телекомунікацій, можуть очікувати більш сильного зростання на так званих «бичачих» ринках. Але коли настає рецесія або «ведмежий» ринок, збитки від цього сектора можуть бути серйозними.

1.2.2 Оцінка міжнародних телекомунікаційних компаній

Важко уникнути висновку, що розмір має значення в телекомунікації. Це дорогий бізнес. Претенденти повинні бути досить великими і виробляти достатній грошовий потік, щоб покрити витрати на розширення мереж і послуг, які, мабуть, застарівають відразу. Системи передачі необхідно замінювати кожні два роки. Великі компанії, які володіють великими мережами, особливо локальними мережами, які простягаються прямо в будинки і офіси клієнтів, менш залежні від взаємодії з іншими компаніями для доставки дзвінків і даних в їх кінцевих пунктів призначення. З іншого боку, дрібні гравці повинні платити за приєднання частіше, щоб закінчити роботу. Для невеликих операторів, які сподіваються вирости великими за один день, фінансові проблеми, пов'язані з швидкими технологічними змінами і зносом обладнання, можуть бути величезними.

Доходи можуть бути складним питанням при аналізі телекомунікаційних компаній. Багато компаній мало заробляють або взагалі не заробляють. Щоб оцінити вартість компанії, аналітики телекомунікаційної галузі можуть звернутися до співвідношення ціни до продажів (ціна акцій, поділена на продажі). Вони також дивляться на середній дохід на користувача (ARPU), який пропонує корисну міру продуктивності зростання, і рівень відтоку, швидкість, з якою клієнти йдуть (імовірно для конкурента).

Поточні лідери галузі в усьому світі можуть змінюватися з року в рік. Визначення того, які з них є найбільшими, залежить від того, з якої точки зору вони виглядають - з точки зору загального обсягу продажів або з точки зору ринкової капіталізації. Станом на 2018 щороку п'ять найбільших телекомунікаційних компаній за ринковою капіталізацією розподілилися наступним чином:

- Verizon (VZ), яка надає послуги бездротового і дротового зв'язку, в додаток до широкосмугових і інформаційних послуг, має поточну ринкову капіталізацію близько 200 мільярдів доларів.

Він залишається привабливим як постачальник дивідендів через свій фінансовий стан: це найбільша телекомунікаційна компанія в США, що працює в 150 країнах.

- China Mobile Ltd (CHL), яка працює в бізнесі тільки з 1997 року, має ринкову капіталізацію приблизно в 185 мільярдів доларів через зростання використання стільникових телефонів і інтернет-послуг в Китаї за останні 15 років.
- AT & T (T), найстаріша компанія в телекомунікаційному бізнесі, має ринкову вартість, яка оцінюється приблизно в 182 мільярди доларів. China Mobile затьмарила AT & T трохи більше ніж за десятиліття своєї діяльності. Проте, на частку колишньої Ma Bell припадає третина частки ринку бездротового зв'язку в США і вражаючи 30-річна історія збільшення дивідендів.
- Vodafone Group (VOD), найбільша телекомунікаційна компанія Сполученого Королівства, надає послуги та обладнання для голосового зв'язку, широкосмугового зв'язку та передачі даних, а її ринкова капіталізація становить близько 97 млрд. Дол. США.
- Японська SoftBank Corp, що надає програмне забезпечення для IT, а також бездротові та інтернет-послуги, має поточну ринкову капіталізацію трохи більше 90 мільярдів доларів.

Рейтинги помітно змінюються, якщо судити по загальному доходу від продажів. У той час як Verizon, China Mobile і AT & T залишаються в п'ятірці лідерів, Verizon опустилася на друге місце після японської корпорації Nippon Telegraph & Telephone Corporation (NTT), яка є великим виробником телекомунікаційного обладнання в додаток до великої телефонної лінії і Інтернету.

1.2.3 Телекомунікаційні ETFs та огляд телекомунікаційного сектора

Кілька біржових фондів (ETF) служать альтернативою прямим інвестиціям в окремі телекомунікаційні фірми. Телекомунікаційні ETF мають різну спрямованість на географію або галузеву спеціалізацію. Деякі з найбільш популярних включають в себе:

- ETF Vanguard Telecommunication Services складається з 98% акцій США, від невеликих регіональних телекомунікаційних компаній до великої трійки, Verizon, AT & T і Sprint.
- Індексний фонд телекомунікаційного сектора США iShares Dow Jones (IYZ), аналогічний за своїми фондами ETF телекомунікаційних послуг Vanguard, також відстежує найбільші телекомунікаційні компанії в США - Sprint, AT & T і Verizon, а також кілька невеликих регіональних постачальників послуг.
- Глобальний телекомунікаційний фонд iShares S & P (IXP) орієнтований на міжнародну діяльність: 70% його акцій перебувають в компаніях зі штаб-квартирою за межами США. Серед основних компаній, що входять в п'ятірку найбільших компаній в області ринкової капіталізації: Verizon, AT & T, China Mobile, Vodafone і Softbank Corp.

Аналітики передбачають, що інновації в продуктах і збільшення злиттів і поглинань будуть тільки сприяти подальшому зростанню і успіху телекомунікаційної галузі. Є багато можливостей для інвесторів, і збільшення інвесторів буде тільки сприяти подальшому розвитку сектора.

3% - середньорічний історичний темп зростання телекомунікаційного сектора.

Стабільність зростання сектора, навіть в періоди спаду, означає, що він вважається надійною захисною інвестицією, зберігаючи при цьому свою привабливість для зростаючих інвесторів. Навіть в нестабільні економічні часи стійкий попит на послуги передачі голосу і даних забезпечує стабільне джерело доходів для великих телекомунікаційних компаній.

Телекомунікації стають все більш важливою базовою галуззю, яка віщує її перспективи на майбутнє і подальше зростання. Постійні досягнення в області високошвидкісних мобільних послуг і підключення до Інтернету між пристроями продовжують стимулювати інновації та конкуренцію в цьому секторі. Велика частина індустрії зосереджена на наданні більш швидких послуг передачі даних, особливо в області відео високої роздільної здатності. По суті, рушійними силами є швидкі і чіткі послуги, розширення можливостей підключення і використання декількох додатків.

Країни з ринковою економікою продовжують залишатися благом для галузі, оскільки темпи зростання індустрії стільникових телефонів в таких країнах, як Китай і Індія, змушують виробників обладнання йти в ногу з рівнем попиту.

У США аналітики приділяють пильну увагу питанням, пов'язаним з нейтральністю мережі, оскільки попит на послуги передачі даних і відео продовжує зростати і в майбутньому. Як і раніше існує високий попит на права бездротового спектра, про що свідчить стимулюючий аукціон FCC, який відбувся в квітні 2017 року, не кажучи вже про зростаючу тенденцію до консолідації за допомогою злиття і поглинань.

1.3 Рейтинг українських інтернет-провайдерів

Французька компанія nPerf, яка спеціалізується на вимірюванні швидкості і якості доступу до мережі інтернет, опублікувала дані тестування мереж українських ШПД-провайдерів. Компанія має в своєму розпорядженні вимірювальні сервери на території України, за допомогою яких за 2018 рік було проведено 437 305 тестів підключення на дев'яти найбільших інтернет-мережах України.

У підсумку, найвища середня швидкість була зафіксована в мережах «Ланет» - 64,82 Мбіт / сек (download) і 66,17 Мбіт / сек (upload). На другому місці розташувався «Тріолан» з показником 54,56 Мбіт / сек і 56,48 Мбіт / сек відповідно, на третьому - «О3 / Freenet» - 48,8 Мбіт / сек і 51,17 Мбіт / сек. Четверте місце зайняв «Київстар» (42,56 Мбіт / сек і 46,67 Мбіт / сек) і замкнула п'ятірку лідерів «Воля» (45,45 Мбіт / сек і 27,58 Мбіт / сек).

Повністю Топ-9 українських провайдерів виглядає наступним чином:

- «Ланет» - 127 693 nPoints (64,82 Мбіт / сек, 66,17 Мбіт / сек, 25,88 мс)
- «Тріолан» - 120 454 nPoints (54,56 Мбіт / сек, 56,48 Мбіт / сек, 30,64 мс)
- «О3 / Freenet» - 118 836 nPoints (48,80 Мбіт / сек, 51,17 Мбіт / сек, 29,22 мс)
- «Київстар» - 113 179 nPoints (42,56 Мбіт / сек, 46,67 Мбіт / сек, 33,80 мс)
- «Воля» - 111 111 nPoints (45,45 Мбіт / сек, 27,58 Мбіт / сек, 32,00 мс)
- «Fregat» - 107 697 nPoints (39,65 Мбіт / сек, 39,48 Мбіт / сек, 39,86 мс)
- «Vega Telecom» - 103 626 nPoints (33,20 Мбіт / сек, 33,58 Мбіт / сек, 39,88 мс)
- «Datagroup» - 95 896 nPoints (31,78 Мбіт / сек, 32,60 Мбіт / сек, 59,67 мс)

- «Укртелеком» - 65 163 nPoints (11,24 Мбіт / сек, 6,10 Мбіт / сек, 65,17 мс)

За кількістю проведених тестів лідирували «Укртелеком», «Київстар» і «Воля»

1.4 Актуальність теми дослідження та визначення існуючих проблем

Телекомунікації як галузь змінилися разом з необхідністю підключення людей - швидше, надійніше і без фізичних обмежень. Ця стійка індустрія перейшла від стаціонарних телефонів і комутованих інтернет-підключень до надання мобільного підключення клієнтам, які постійно знаходяться в дорозі.

Телекомунікаційні компанії, які прагнуть залишатися конкурентоспроможними на цьому мінливому ринку, повинні з'ясувати, як постійно надавати продукти і послуги, які їм потрібні, і за ціною, що дозволяє їм отримувати прибуток, необхідну для того, щоб залишатися на плаву. Це тонкий баланс, який вимагає як розуміння того, що насправді хвилює клієнтів, так і технологічних можливостей для його забезпечення.

Ось п'ять основних завдань, що стоять сьогодні перед індустрією зв'язку.

№1: зростаючі обсяги, скорочення доходів

Збільшення споживчого попиту зазвичай означає більший дохід, вірно? Не завжди. Як недавно сказав один директор з телекомунікацій на міжнародній галузевій конференції, телекомунікації є «єдиною галуззю, в якій обсяги зростають на 40-60 відсотків на рік, але доходи скорочуються».

Згідно International Telecommunication Union News, це означає, що телекомунікаційні компанії, які прагнуть «повернути назад тенденцію до

зниження доходів і підключити більше клієнтів до більшої кількості нових послуг по всьому світу», повинні:

- Змінити свої бізнес-моделі
- Зробити капітальний ремонт експлуатаційних процедур
- Знову провести навчання для департаментів підбору персоналу та «людських ресурсів»

На щастя, телекомунікації вже використовують корисні ідеї та відповіді в мільярдах рядків даних, що зберігаються, які вони неминуче накопичили. Новітня аналітика комунікацій допомагає організаціям перетворювати структуровані дані в ідеї, на які вони можуть впливати. Використовуючи існуючі інструменти, особи, які приймають рішення в телекомунікаційних організаціях, можуть отримати миттєві відповіді, а потім застосувати цю інформацію в дії. В даний час найкраща бізнес-аналітика означає кращі бізнес-результати.

№2: Очікування зниження ризику кібербезпеки

Телекомунікації продовжують боротися з ризиками кібербезпеки мережі. Клієнти очікують, що їх оператори збережуть їх особисту інформацію та мережі в безпеці. Оскільки нові загрози постійно знаходяться на горизонті, це означає, що компанії повинні постійно виявляти пильність у боротьбі з витоками даних і зламами даних. Утримання нового бізнесу і утримання існуючих клієнтів все частіше стануть справою демонстрації прихильності зниження ризику кібербезпеки і припинення атак до того, як вони відбудуться, щоб компанії не відчували підірваної довіри клієнтів після успішної атаки.

№3: Швидке розширення Інтернету речей (IoT)

IoT - це захоплююча телекомунікація, яку продовжать вивчати та розробляти. Той факт, що багато пристроїв стають здатними підключатися до мереж передачі даних, вражає. Це значна можливість для заповзятливих

операторів зв'язку, які прагнуть розширити свої пропозиції і врахувати прагнення людей до розширення взаємодії. Але це також завдання і серйозне зобов'язання, щоб зрозуміти це правильно і з'ясувати, як монетизувати ці зусилля.

Коли мова заходить про роль телекомунікаційних компаній в розширенні IoT, надійність і конфіденційність стануть одними з найбільш насущних проблем.

№ 4: Зростаюча потреба в диверсифікації пропозицій

Отримання прибутку в конкурентній телекомунікаційної галузі потребують все більшої диверсифікації. Самовдоволення може означати фінансову смерть для вашого прибутку. Для телекомунікаційних компаній стане все більш важливим розширювати свої основні пропозиції і виходити на нові ринки.

№ 5: Попит на чуйне обслуговування клієнтів

Чуйне обслуговування клієнтів буде продовжувати псувати репутацію і результати компаній. Те, як телекомунікаційні компанії справляються з перебоями в обслуговуванні і скаргами клієнтів, буде грати ще більшу роль в лояльності клієнтів або їх відсутності. Люди очікують своєчасної відповіді, незалежно від того, зазначають вони компанію в повідомленні в соціальній мережі або беруть участь в живому чаті з фахівцем з обслуговування клієнтів. Скорочення відтоку клієнтів вимагає від компаній пропонувати доступне і корисне обслуговування клієнтів в будь-яких ситуаціях.

Ці п'ять проблем, що стоять перед галуззю зв'язку, є прихованими можливостями для компаній, що прагнуть швидко розвиватися.

1.5 Висновки до 1 розділу

У цьому розділі розглянуто теоретичні відомості телекомунікаційної сфери, еволюцію телекомунікаційного сектора, дослідження нинішньої ситуації на телекомунікаційному ринку України. Також виконано обґрунтування актуальності дослідження, особливості розвитку процесів у телекомунікаціях на сучасному етапі, рейтинг українських Інтернет-провайдерів.

РОЗДІЛ 2 МЕТОДИ МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ ПРОЦЕСІВ НА ТЕЛЕКОМУНІКАЦІЙНОМУ РИНКУ УКРАЇНИ

2.1 Дослідження телекомунікаційного ринку України. Підходи до аналізу

Частка ринку є іншим важливим показником маркетингових зусиль підприємства і характеризує стан підприємства щодо конкурентів. Кількісно частка ринку визначається як відношення обсягу продажу товарів підприємства до загального обсягу продажу товарів тієї ж категорії на ринку, виражене у відсотках.

Частка ринку може бути розрахована різними способами:

- за обсягом продажу товарів в натуральному вираженні 1.1:

$$D_p = \frac{Q_i}{E_p} \quad (1.1)$$

де Q_i - обсяг продажів, який може реалізувати фірма; E_p - ємність ринку

- за формулою 1.2:

$$D_p = \frac{Q_i}{Q_{max}} \quad (1.2)$$

де Q_{max} - фірма, що має максимальний обсяг продажів (у вартісному вираженні; в обслуговується сегменті; щодо продажу ряду найближчих конкурентів; щодо лідера ринку, провідного конкурента).

Абсолютна частка ринку D_i фірми і на ринку визначається за формулою 1.3, як відношення товарообороту фірми T_i до загального обсягу продажів на ринку ($\sum T_i$):

$$D_i = \frac{T_i}{\sum T_i} \quad (1.3)$$

Відносна частка ринку розраховується за формулою 1.4:

$$D_i = \frac{T_i}{\sum T_{ki}} \quad (1.4)$$

де T_{ki} - обсяг продажів основних конкурентів.

Метод ланцюгових відносин може бути використаний як для оцінки місткості ринку товарів виробничого призначення, так і для товарів широкого споживання. Суть методу полягає в наступному.

1. Будується робоча гіпотеза про залежність ємності ринку від ряду ринкових чинників. При цьому передбачається, що зазначена залежність визначається за формулою 1.5:

$$E = K_1 \cdot K_2 \cdot K_3 \cdot \dots \cdot K_n \quad (1.5)$$

де K_1 , K_2 і т.д. - коефіцієнти, що відображають вплив кожного з ринкових факторів на ємність ринку. При цьому кожен наступний (зліва направо) коефіцієнт уточнює результат, отриманий за допомогою введення попереднього коефіцієнта.

2. За допомогою кабінетних або польових досліджень розраховуються чисельні значення коефіцієнтів K_1 , K_2 , K_n .

3. Обчислюється ємність ринку.

До якісних оцінок ринку можна віднести наступні:

1. Рейтинги марок на ринку (за рейтингом в аналізі рекламних ЗМІ, частоті, площі, часу реклами; за даними експертів).

2. Відповідність частки ринку (велика, середня, мала), яку можна встановити виходячи з АВС-аналізу (гіпотеза Парето): велика частка (група А), до неї відносяться 15% великих марок, торгових фірм або виробників, що забезпечують 65% обороту ринку; середня (група В), коли наступні 20%

марок забезпечують 20% решти ринку; мала - С, коли 65%, що залишилися марок мають ринкову частку в 15%.

3. Рівень якості послуг, що надаються за даними експертних оцінок. На підставі сукупних відгуків споживачів і конкурентів можна досить об'єктивно оцінити рівень якості послуг конкретної фірми або марки. Доцільно використовувати комплексний показник якості обслуговування, який вираховується за формулою 1.6:

$$K_{\text{кач}} = \frac{\sum x_1 + \sum x_2 - \sum x_3}{\sum x_1 + \sum x_2 + \sum x_3 + \sum x_4} \quad (1.6)$$

де $\sum x_1$ - число респондентів, які оцінили якість послуг, що надаються на "відмінне"; $\sum x_2$ - "добре"; $\sum x_3$ - "задовільний"; $\sum x_4$ - "незадовільний". Сума значень показників обслуговування визначиться з опитувальних таблиць.

Чим вище значимість $K_{\text{кач}}$, тим краще обслуговування. Межа значущості може бути дорівнює 1, коли всі покупці оцінили якість обслуговування як "відмінне".

Різновидом оцінки виступає наступна методика. При оцінці якості обслуговування використовують узагальнений показник До оцінок чотирьох груп покупців - А, В, С, D, основу якого складають відгуки покупців за формулою 1.7 :

$$K_o = \frac{X_1 \cdot 5 + X_2 \cdot 4 + X_3 \cdot 3 + X_4 \cdot 2}{X_1 + X_2 + X_3 + X_4} \quad (1.7)$$

де X_1 - кількість респондентів, які привласнили оцінку "відмінно"; X_2 - кількість з оцінкою "добре"; X_3 - "задовільно"; X_4 - "незадовільно". Діапазон значущості прагне до п'яти, коли всі покупці оцінили обслуговування як відмінне і гарне. Чим вище значимість коефіцієнта K (), тим краще обслуговування.

Статистичний аналіз експертних оцінок.

За даними проведеного експертного опитування (метод Дельфі) про перспективи обсягу продажів товару X складається таблиця опитування експертів.

Величина обсягу продажів, за яку висловилися найбільша кількість експертів (мода), характеризує варіант обсягу продажів, який найбільш часто зустрічається в даному варіаційному ряду. Цей показник при прогнозуванні обсягу продажів має велике значення, так як з його допомогою можна виявити переважне судження фахівців з досліджуваного питання.

2.2 Аналітична частина

Публічне акціонерне товариство «Укртелеком» — найбільший оператор фіксованого зв'язку в Україні. Компанія надає широкий спектр телекомунікаційних послуг в усіх регіонах країни. Уже багато років її справедливо вважають лідером серед провайдерів фіксованого інтернету.

Укртелеком має найбільше інтернет-покриття та найрозгалуженішу оптичну мережу.

У 2011 році компанія заснувала оператора «ТриМоб», який надає послуги мобільного зв'язку за технологією UMTS/WCDMA.

У 2015 році Укртелеком розпочав масштабну модернізацію мережі.

Укртелеком пропонує практично всі види сучасних телекомунікаційних послуг.

Для домашніх користувачів:

- Телефонія;
- Інтернет;
- Інтерактивне TV

- Для бізнес-клієнтів:
- Оптичний інтернет
- Передача даних та інтернет
- Інтернет-телефонія
- Телефонія
- Хмарна АТС
- Дата-центр
- Організації віртуальних приватних мереж (VPN)
- Номери для організації «гарячих» ліній
- Зарубіжні хмарні сервіси
- Захист від DDoS атак
- TV Бізнес

На рисунках (2.1-2.4) нижче представлені данні за 2019, на яких можна побачити кількість підключень та відключень фіксованого зв'язку та кількість підключень та відключень інтернету в системі B2C та B2B.

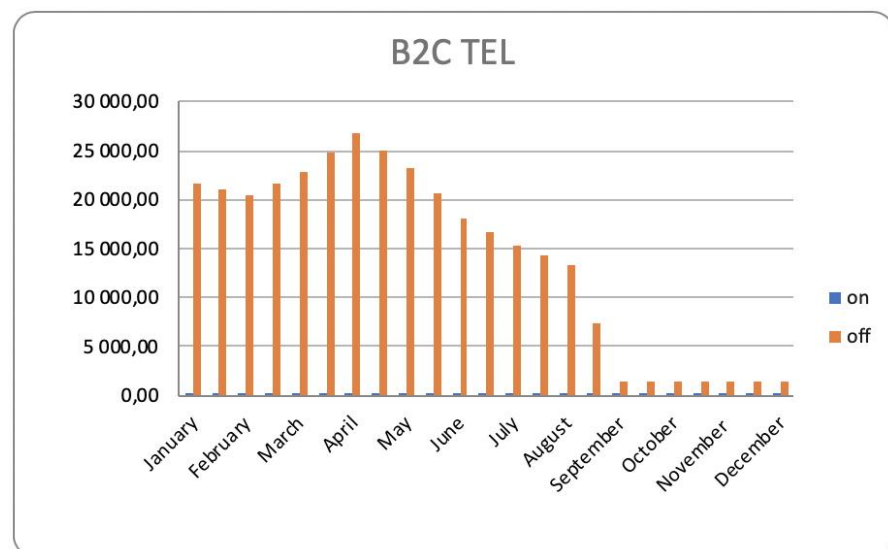


Рис. 2.1 – кількість підключень та відключень фіксованого зв'язку в сегменті B2C

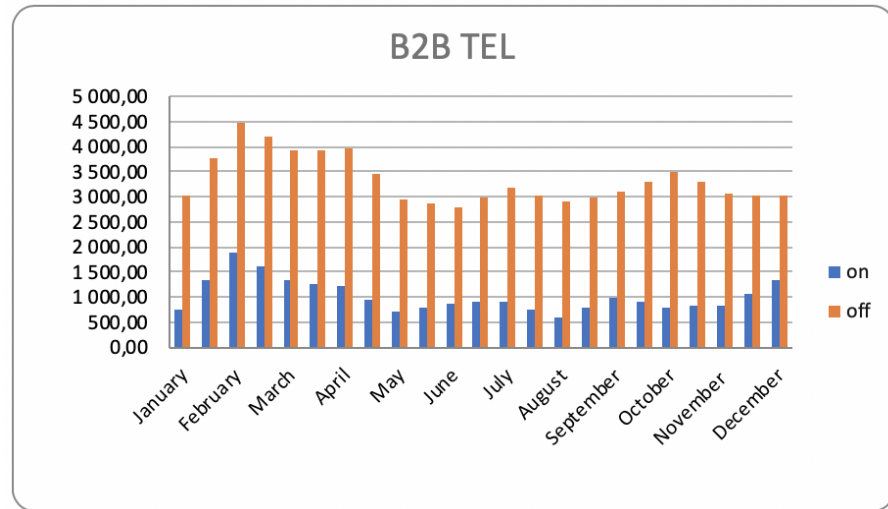


Рис. 2.2 – кількість підключень та відключень фіксованого зв'язку в сегменті B2B

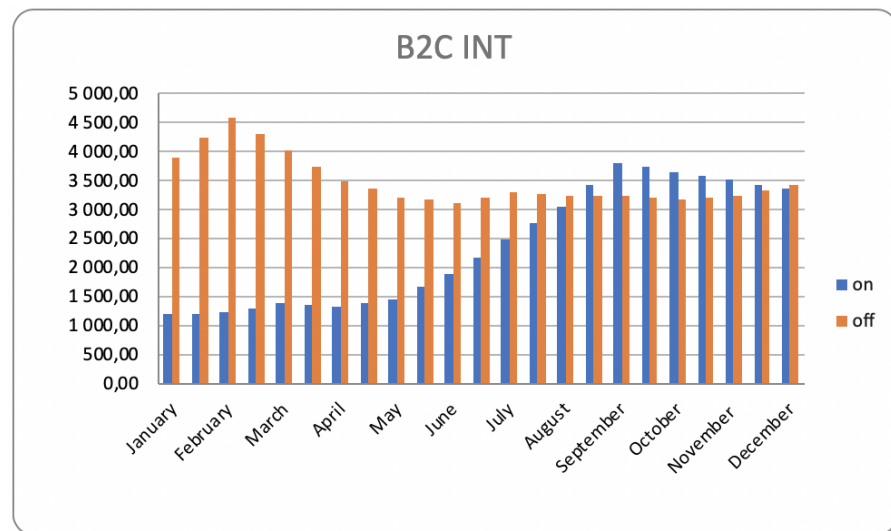


Рис. 2.3 – кількість підключень та відключень інтернету в сегменті B2C

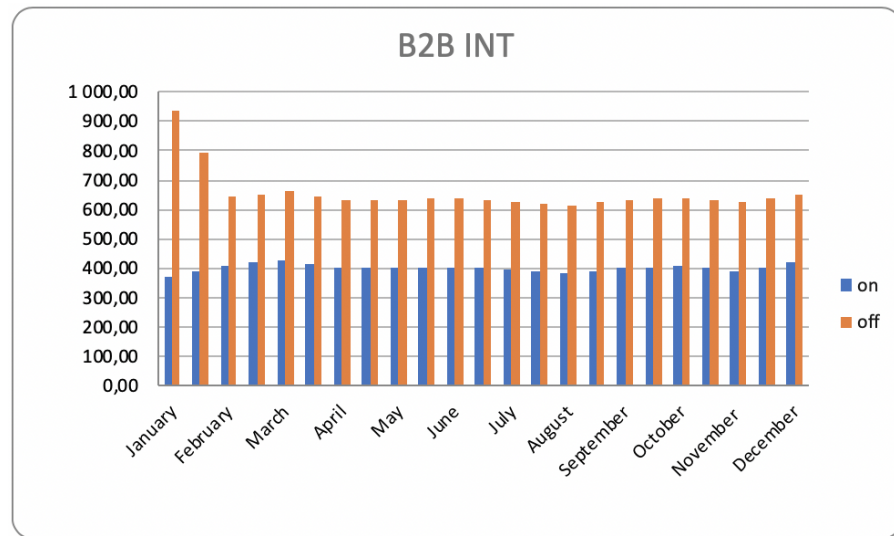


Рис. 2.4 – кількість підключень та відключень інтернету в сегменті B2B

Нижче приведена розмноження вибірки даних та таблиці даних.

Якщо є 12 чисел (помісячно), то знаходимо середнє для кожної пари і отримаємо:

$$11+12 = 23 \text{ числа}$$

Часовий ряд складається з 23 чисел.

Таблиці даних: представлені в таблиці 1.

Табл. 1 – таблиця даних ПАТ «Укртелеком» за 2019 рік

| Місяць | B2C | | | | B2B | | | |
|----------|--------------------|-----------|----------|----------|--------------------|----------|----------|--------|
| | Фіксований зв'язок | | Інтернет | | Фіксований зв'язок | | Інтернет | |
| | on | off | on | off | on | off | on | off |
| January | 290.00 | 21,704.00 | 1,183.00 | 3,901.00 | 763.00 | 3,047.00 | 373.00 | 938.00 |
| | 275.00 | 21,093.00 | 1,199.00 | 4,240.50 | 1,337.00 | 3,755.00 | 392.00 | 790.50 |
| February | 260.00 | 20,482.00 | 1,215.00 | 4,580.00 | 1,911.00 | 4,463.00 | 411.00 | 643.00 |
| | 245.00 | 21,666.50 | 1,303.50 | 4,295.00 | 1,626.00 | 4,190.50 | 420.50 | 652.50 |
| March | 230.00 | 22,851.00 | 1,392.00 | 4,010.00 | 1,341.00 | 3,918.00 | 430.00 | 662.00 |
| | 240.00 | 24,787.00 | 1,359.50 | 3,746.50 | 1,277.00 | 3,940.00 | 415.50 | 647.50 |
| April | 250.00 | 26,723.00 | 1,327.00 | 3,483.00 | 1,213.00 | 3,962.00 | 401.00 | 633.00 |

| | | | | | | | | |
|-----------|--------|-----------|----------|----------|----------|----------|--------|--------|
| | 255.00 | 24,951.00 | 1,393.00 | 3,347.00 | 957.00 | 3,448.50 | 402.00 | 634.00 |
| May | 260.00 | 23,179.00 | 1,459.00 | 3,211.00 | 701.00 | 2,935.00 | 403.00 | 635.00 |
| | 205.00 | 20,633.00 | 1,665.00 | 3,159.50 | 783.00 | 2,859.00 | 404.00 | 636.00 |
| June | 150.00 | 18,087.00 | 1,871.00 | 3,108.00 | 865.00 | 2,783.00 | 405.00 | 637.00 |
| | 190.00 | 16,729.50 | 2,174.00 | 3,201.00 | 895.00 | 2,976.50 | 400.00 | 632.00 |
| July | 230.00 | 15,372.00 | 2,477.00 | 3,294.00 | 925.00 | 3,170.00 | 395.00 | 627.00 |
| | 205.00 | 14,346.50 | 2,757.00 | 3,272.00 | 752.00 | 3,032.00 | 388.00 | 620.00 |
| August | 180.00 | 13,321.00 | 3,037.00 | 3,250.00 | 579.00 | 2,894.00 | 381.00 | 613.00 |
| | 190.00 | 7,330.50 | 3,423.00 | 3,245.50 | 779.50 | 3,007.00 | 391.00 | 623.00 |
| September | 200.00 | 1,340.00 | 3,809.00 | 3,241.00 | 980.00 | 3,120.00 | 401.00 | 633.00 |
| | 200.00 | 1,380.00 | 3,722.50 | 3,214.00 | 895.00 | 3,311.00 | 404.00 | 636.00 |
| October | 200.00 | 1,420.00 | 3,636.00 | 3,187.00 | 810.00 | 3,502.00 | 407.00 | 639.00 |
| | 200.00 | 1,385.00 | 3,574.00 | 3,218.50 | 827.00 | 3,285.00 | 399.50 | 631.50 |
| November | 200.00 | 1,350.00 | 3,512.00 | 3,250.00 | 844.00 | 3,068.00 | 392.00 | 624.00 |
| | 200.00 | 1,370.00 | 3,432.00 | 3,337.00 | 1,084.00 | 3,042.00 | 405.00 | 637.00 |
| December | 200.00 | 1,390.00 | 3,352.00 | 3,424.00 | 1,324.00 | 3,016.00 | 418.00 | 650.00 |

2.3 Існуючі моделі прогнозування

До недавнього часу (середини 80-х років минулого століття) існувало кілька загальновизнаних методів прогнозування часових рядів:

- економетричні
- регресивні
- Методи Боксу-Дженкінса (ARIMA, ARMA)

Однак, починаючи з кінця 80-х років, в науковій літературі був опублікований ряд статей по нейросетевой тематиці, в яких був приведений ефективний алгоритм навчання нейронних мереж і доведена можливість їх використання для самого широкого кола завдань.

Ці статті відродили інтерес до нейромереж в науковому співтоваристві і останні дуже скоро стали широко використовуватися при дослідженнях в самих різних областях науки від експериментальної фізики і хімії до економіки.

До речі, деякі співробітники компанії "Нейропроект" почали займатися нейронними мережами саме в той час, використовуючи їх для обробки результатів фізичних експериментів.

Частково через відносну складність і недетермінованості нейронних мереж і генетичних алгоритмів, ці технології не відразу вийшли за рамки суто наукового застосування. Проте, з плином часу рівень довіри до нових технологій підвищувався і з боку бізнесу. З початку 90-х років почали регулярно з'являтися повідомлення про установках нейромережових систем в різних компаніях, банках, корпоративних інститутах. Причому сфера використання нових технологій була дуже багатогранною - оцінка ризиків, контроль технологічних процесів, управління роботами і багато іншого. Одним з найбільш успішних програм нейронних мереж було прогнозування

часових рядів. Причому наймасовішим було прогнозування на фінансових ринках і прогнозування продажів

В даний час можна з упевненістю сказати, що використання нейронних мереж при прогнозуванні дає відчутну перевагу в порівнянні з більш простими статистичними методами.

Методи прогнозування, засновані на згладжування, експоненційному згладжуванні і ковзному середньому

"Наївні" моделі прогнозування

При створенні "наївних" моделей передбачається, що деякий останній період прогнозованого тимчасового ряду найкраще описує майбутнє цього прогнозованого ряду, тому в цих моделях прогноз, як правило, є дуже простою функцією від значень прогнозованої змінної в недалекому минулому.

Найпростішою моделлю є $Y(t + 1) = Y(t)$

що відповідає припущенням, що "завтра буде як сьогодні".

Поза всяким сумнівом, від такої примітивної моделі не варто чекати великої точності. Вона не тільки не враховує механізми, що визначають прогнозовані дані (цей серйозний недолік взагалі властивий багатьом статистичним методам прогнозування), але і не захищена від випадкових флуктуацій, вона не враховує сезонні коливання і тренди. Втім, можна будувати "наївні" моделі дещо по-іншому.

$$Y(t + 1) = Y(t) + [Y(t) - Y(t - 1)],$$

$$Y(t + 1) = Y(t) \cdot \left[\frac{Y(t)}{Y(t - 1)} \right]$$

такими способами ми намагаємося пристосувати модель до можливих трендам $Y(t + 1) = Y(t - s)$ це спроба врахувати сезонні коливання

Середні і ковзаючі середні.

Найпростішою моделлю, заснованою на простому усередненні є $Y(t + 1) = \left(\frac{1}{t}\right) \cdot [Y(t) + Y(t - 1) + \dots + Y(1)]$, і на відміну від найпростішої "наївної" моделі, якій відповідав принцип "завтра буде як сьогодні", цієї моделі відповідає принцип "завтра буде як було в середньому за останній час". Така модель, звичайно стійкіша до флуктуацій, оскільки в ній згладжуються випадкові викиди щодо середнього. Незважаючи на це, цей метод ідеологічно настільки ж примітивний як і "наївні" моделі і йому властиві майже ті ж самі недоліки.

У наведеній вище формулі передбачалося, що ряд усереднюється по досить тривалому інтервалу часу. Однак як правило, значення часового ряду з недалекого минулого краще описують прогноз, ніж старіші значення цього ж ряду. Тоді можна використовувати для прогнозування ковзне середнє $Y(t + 1) = \left(\frac{1}{T+1}\right) \cdot [Y(t) + Y(t - 1) + \dots + Y(t - T)]$.

Сенс його полягає в тому, що модель бачить тільки найближче минуле (на T відліків за часом в глибину) і ґрунтуючись тільки на цих даних будує прогноз.

При прогнозуванні досить часто використовується метод експоненційних середніх, який постійно адаптується до даних за рахунок нових значень. Формула 1.8, що описує цю модель записується як:

$$Y(t + 1) = a \cdot Y(t) + (1 - a)^{Y(t)}, \quad (1.8)$$

де $Y(t + 1)$ - прогноз на наступний період часу $Y(t)$ - реальне значення в момент часу $t^{Y(t)}$ - минулий прогноз на момент часу t , a - постійна згладжування ($0 \Leftarrow a \Leftarrow 1$)

У цьому методі є внутрішній параметр a , який визначає залежність прогнозу від старіших даних, причому вплив даних на прогноз експоненціально убуває з "віком" даних. Залежність впливу даних на прогноз при різних коефіцієнтах a наведена на графіку.

Видно, що при $a \rightarrow 1$, експоненціальна модель прагне до найпростішої "наївної" моделі. При $a \rightarrow 0$, прогнозована величина стає рівною попередньому прогнозу.

Якщо проводиться прогнозування з використанням моделі експоненціального згладжування, зазвичай на деякому тестовому наборі будуються прогнози при $a = [0.01, 0.02, \dots, 0.98, 0.99]$ і відстежується, при якому a точність прогнозування вище. Це значення a потім використовується при прогнозуванні надалі.

Хоча описані вище моделі ("наївні" алгоритми, методи, засновані на середніх, що ковзають середніх і експоненціального згладжування) використовуються при бізнес-прогнозуванні в не дуже складних ситуаціях, наприклад, при прогнозуванні продажів на спокійних і усталених західних ринках, ми не рекомендуємо використовувати ці методи в задачах прогнозування на увазі явної примітивності і неадекватності моделей.

Разом з цим хотілося б зазначити, що описані алгоритми цілком успішно можна використовувати як супутні і допоміжні для предоброботки даних в задачах прогнозування. Наприклад, для прогнозування продажів в більшості випадків необхідно проводити декомпозицію тимчасових рядів (тобто виділяти окремо тренд, сезонну і нерегулярну складові). Одним з методів виділення трендових складових є використання експоненціального згладжування.

Методи Хольта і Брауна

В середині минулого століття Хольт запропонував вдосконалений метод експоненціального згладжування, згодом названий його ім'ям. У запропонованому алгоритмі значення рівня та тренда згладжуються за допомогою експоненціального згладжування. Причому параметри згладжування у них різні.

$$\Omega_t = \alpha Y_t + (1 - \alpha)(\Omega_{t-1} - T_{t-1}), \quad (1.9)$$

$$T_t = \beta(\Omega_t - \Omega_{t-1}) + (1 - \beta)T_{t-1}, \quad (1.10)$$

$$Y_{t+p} = \Omega_t + p \cdot T_t \quad (1.11)$$

Тут рівняння 1.9 описує згладжений ряд загального рівня.

Рівняння 1.10 служить для оцінки тренда.

Рівняння 1.11 визначає прогноз на p відліків за часом вперед.

Постійні згладжування в методі Хольта ідеологічно грають ту ж роль, що і постійна в простому експоненціальному згладжуванні. Підбираються вони, наприклад, шляхом перебору по цих параметрах з якимсь кроком. Можна використовувати і менш складні в сенсі кількості обчислень алгоритми. Головне, що завжди можна підібрати таку пару параметрів, яка дає велику точність моделі на тестовому наборі і потім використовувати цю пару параметрів при реальному прогнозуванні.

Окремим випадком методу Хольта є метод Брауна, коли $\alpha = 0$.

Метод Вінтерса

Хоча описаний вище метод Хольта (метод двопараметричного експоненціального згладжування) і не є зовсім простим (щодо "наївних" моделей і моделей, заснованих на усередненні), він не дозволяє враховувати сезонні коливання при прогнозуванні. Говорячи більш акуратно, цей метод не може їх "бачити" в передісторії. Існує розширення методу Хольта до трипараметричного експоненціального згладжування. Цей алгоритм називається методом Вінтерса. При цьому робиться спроба врахувати сезонні складові в даних. Система рівнянь 1.12, що описують метод Вінтерса виглядає наступним чином:

$$\begin{cases} \Omega_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(\Omega_{t-1} - T_{t-1}), \\ T_t = \beta(\Omega_t - \Omega_{t-1}) + (1 - \beta)T_{t-1}, \\ S_t = \gamma \frac{Y_t}{\Omega_t} + (1 - \gamma)S_{t-s}, \\ \hat{Y}_{t+p} = (\Omega_t + pT_t)S_{t-s+p} \end{cases} \quad (1.12)$$

Дріб в першому рівнянні служить для виключення сезонності з $Y(t)$. Після виключення сезонності алгоритм працює з "чистими" даними, в яких немає сезонних коливань. З'являються вони вже в самому фінальному прогнозі, коли "чистий" прогноз, порахований майже по методу Хольта множиться на сезонний коефіцієнт.

Регресивні методи прогнозування

Поряд з описаними вище методами, заснованими на експоненціальному згладжуванні, вже досить довгий час для прогнозування використовуються регресійні алгоритми. Коротко суть алгоритмів такого класу можна описати так.

Існує прогнозована змінна Y (залежна змінна) і відібраний заздалегідь комплект змінних, від яких вона залежить - X_1, X_2, \dots, X_n (незалежні змінні). Природа незалежних змінних може бути різною. Наприклад, якщо припустити, що Y - рівень попиту на деякий продукт в наступному місяці, то незалежними змінними можуть бути рівень попиту на цей же продукт в минулий і позаминулий місяці, витрати на рекламу, рівень платоспроможності населення, економічна обстановка, діяльність конкурентів і багато іншого. Головне - вміти формалізувати всі зовнішні чинники, від яких може залежати рівень попиту в числову форму.

Модель множинної регресії в загальному випадку описується виразом 1.13:

$$Y = \mathcal{F}(X_1, X_2, \dots, X_n) + \varepsilon \quad (1.13)$$

У більш простому варіанті лінійної регресійної моделі залежність залежної змінної від незалежних має вигляд 1.14:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (1.14)$$

Тут $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ - підбираються коефіцієнти регресії,

ε - компонента помилки. Передбачається, що всі помилки незалежні і нормально розподілені.

За допомогою таблиці значень минулих спостережень можна підібрати (наприклад, методом найменших квадратів) коефіцієнти регресії, налаштувавши тим самим модель.

При роботі з регресією треба дотримуватися певної обережності і обов'язково перевірити на адекватність знайдені моделі. Існують різні способи такої перевірки. Обов'язковою є статистичний аналіз залишків, тест Дарбіна-Уотсона. Корисно, як і в випадку з нейронними мережами, мати незалежний набір прикладів, на яких можна перевірити якість роботи моделі.

Методи Боксу-Дженкінса (ARIMA)

В середині 90-х років минулого століття був розроблений принципово новий і досить потужний клас алгоритмів для прогнозування часових рядів. Велику частину роботи по дослідженню методології та перевірці моделей була проведена двома статистиками, Г.Е.П. Боксом (G.E.P. Box) і Г.М. Дженкінсом (G.M. Jenkins). З тих пір побудова подібних моделей і отримання на їх основі прогнозів іноді називається методами Боксу-Дженкінса. Більш докладно ієрархію алгоритмів Боксу-Дженкінса ми розглянемо трохи нижче, поки ж відзначимо, що в це сімейство входить кілька алгоритмів, найвідомішим і використовуваним з них є алгоритм ARIMA. Він вбудований практично в будь-який спеціалізований пакет для прогнозування. У класичному варіанті ARIMA не використовуються незалежні змінні. Моделі

спираються тільки на інформацію, що міститься в передісторії прогнозованих рядів, що обмежує можливості алгоритму. В даний час в науковій літературі часто згадуються варіанти моделей ARIMA, що дозволяють враховувати незалежні змінні. В даному підручнику ми їх розглядати не будемо, обмежившись тільки загальновідомим класичним варіантом. На відміну від розглянутих раніше методик прогнозування тимчасових рядів, в методології ARIMA передбачає будь-якої чіткої моделі для прогнозування даної тимчасової серії. Здається лише загальний клас моделей, що описують часовий ряд і дозволяють якось виражати поточне значення змінної через її попередні значення. Потім алгоритм, підлаштовуючи внутрішні параметри, сам вибирає найбільш підходящу модель прогнозування. Як вже зазначалося вище, існує ціла ієрархія моделей Боксу-Дженкінса. Логічно її можна визначити так

$$AR(p) + MA(q) \rightarrow ARMA(p, q) (P, Q) \rightarrow ARIMA(p, q, r) (P, Q, R) \rightarrow \dots$$

$AR(p)$ - авторегресійна модель порядку p . Модель має вигляд: $Y(t) = f_0 + f_1 \cdot Y(t-1) + f_2 \cdot Y(t-2) + \dots + f_p \cdot Y(t_p) + E(t)$, де $Y(t)$ - залежна змінна в момент часу t . $f_0, f_1, f_2, \dots, f_p$ - оцінювані параметри. $E(t)$ - помилка від впливу змінних, які не враховуються в цій моделі. Завдання полягає в тому, щоб визначити $f_0, f_1, f_2, \dots, f_p$. Їх можна оцінити різними способами. Правильніше за все шукати їх через систему рівнянь Юла-Уолкера, для складання цієї системи буде потрібно розрахунок значень автокореляційної функції. Можна вчинити більш простим способом - порахувати їх методом найменших квадратів.

$MA(q)$ - модель з ковзаючим середнім порядку q . Модель має вигляд 1.15:

$$Y(t) = m + e(t) - w_1 \cdot e(t-1) - w_2 \cdot e(t-2) - \dots - w_p \cdot e(t_p) \quad (1.15)$$

Де $Y(t)$ - залежна змінна в момент часу t . $w_0, w_1, w_2, \dots, w_p$ - оцінювані параметри.

2.4 Програмне забезпечення для прогнозування, переваги та недоліки

EViews – це статистичний пакет для Windows, який використовується, головним чином, для економетричного аналізу, орієнтованого на часові ряди.

EViews застосовується для фінансового аналізу, прогнозування стану ринків, макроекономічного прогнозування та іншого.

Переваги:

- Величезний набір сучасних методів для просунутої економетрики.
- Докладна довідкова система.
- Легкий командний синтаксис і інтерфейс.
- Легка відтворюваність моделей і отримання графіків.
- Можливість створення власних макросів.

Недоліки:

- Невеликі можливості візуалізації.

Побудова мат. моделі:

Побудуємо математичні моделі в Eviews.

1. Побудова авторегресії $AR(p)$
2. $AR(1)$
3. $AR + trend$

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + b_1 k + b_2 k^2 + \dots + b_7 k^7 + \varepsilon(k)$$

Потім викидаємо з моделі статистично незначущі b_i - коефіцієнти моделі.

Нижче приведені отримані результати роботи ПЗ Eviews 3.0 з існуючими даними.

1. Підключення фіксованого зв'язку в сегменті B2C

Dependent Variable: Y

Method: Least Squares

Sample(adjusted): 5 23

Included observations: 19 after adjusting endpoints

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4)$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C(1) | 84.33349 | 41.29890 | 2.042028 | 0.0591 |
| C(2) | 0.777942 | 0.211647 | 3.675657 | 0.0022 |
| C(3) | -0.455468 | 0.200910 | -2.267027 | 0.0386 |
| C(4) | 0.266152 | 0.135292 | 1.967241 | 0.0679 |
| R-squared | 0.570387 | Mean dependent var | 211.6316 | |
| Adjusted R-squared | 0.484464 | S.D. dependent var | 27.25642 | |
| S.E. of regression | 19.57033 | Akaike info criterion | 8.970571 | |
| Sum squared resid | 5744.970 | Schwarz criterion | 9.169400 | |
| Log likelihood | -81.22042 | Durbin-Watson stat | 1.730767 | |

Побудова корелограми зображено на рисунку 2.5

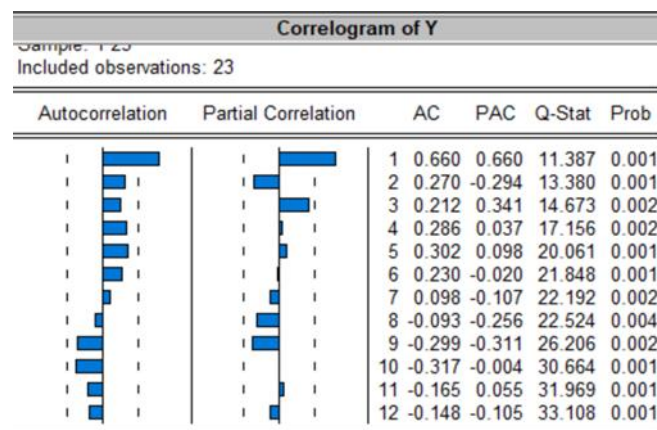


Рис. 2.5 – корелограма підключення фіксованого зв'язку B2C

Побудова прогнозування зображено на рисунку 2.6

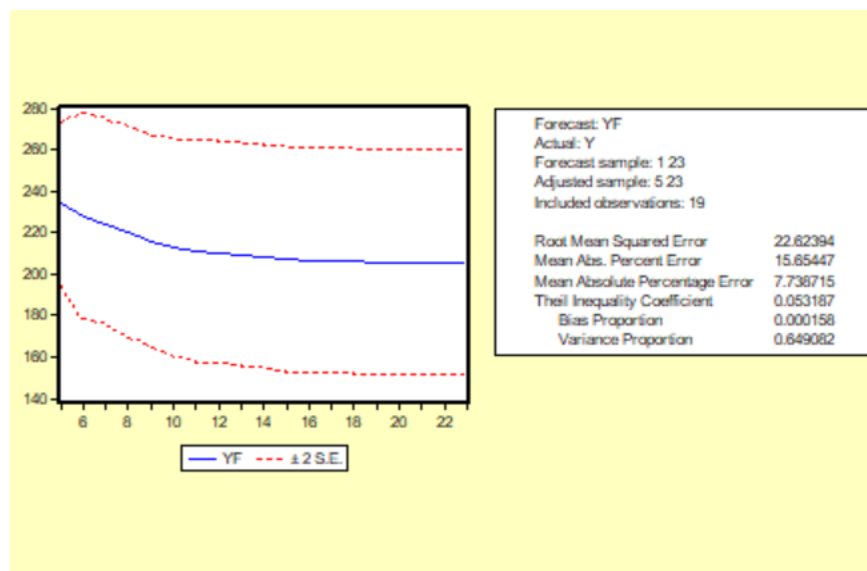


Рис. 2.6 – прогнозування підключення фіксованого зв'язку В2С

2. Відключення фіксованого зв'язку в сегменті В2С

Dependent Variable: Y

Method: Least Squares

Sample(adjusted): 5 23

Included observations: 19 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4)$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|------------|-----------------------|----------|
| C(1) | -199.0022 | 413.3267 | -0.481465 | 0.6371 |
| C(2) | 1.587578 | 0.281895 | 5.631804 | 0.0000 |
| C(3) | -0.689325 | 0.279251 | -2.468476 | 0.0261 |
| C(4) | 0.072315 | 0.063392 | 1.140765 | 0.2719 |
| R-squared | 0.968678 | | Mean dependent var. | 12523.42 |
| Adjusted R-squared | 0.962414 | | S.D. dependent var | 9859.871 |
| S.E. of regression | 1911.540 | | Akaike info criterion | 18.13387 |

| | | | |
|-------------------|-----------|--------------------|----------|
| Sum squared resid | 54809785 | Schwarz criterion | 18.33270 |
| Log likelihood | -168.2718 | Durbin-Watson stat | 1.873076 |

Побудова корелограми зображено на рисунку 2.7

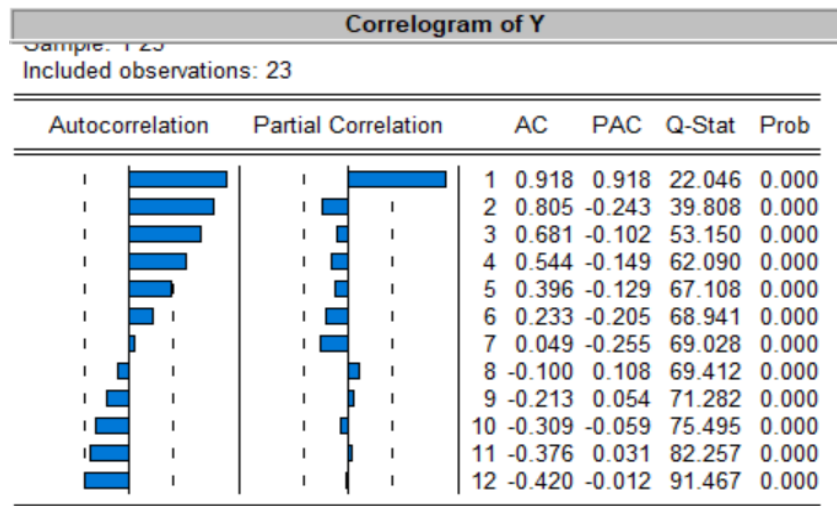


Рис. 2.7 - корелограма відключення фіксованого зв'язку В2С

Побудова прогнозування зображено на рисунку 2.8

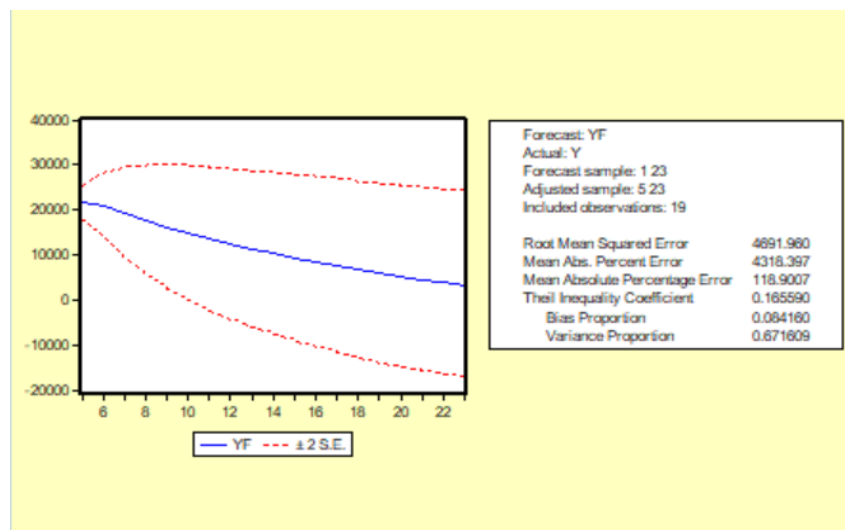


Рис. 2.8 - прогнозування відключення фіксованого зв'язку В2С

3. Підключення інтернету в сегменті В2С

Dependent Variable: Y

Method: Least Squares

Sample(adjusted): 5 23

Included observations: 19 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4)$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|------|-------------|------------|-------------|--------|
| C(1) | 769.9301 | 299.9090 | 2.567212 | 0.0215 |
| C(2) | 1.151488 | 0.321535 | 3.581224 | 0.0027 |
| C(3) | -0.314847 | 0.274510 | -1.146944 | 0.2694 |
| C(4) | -0.069300 | 0.050286 | -1.378122 | 0.1884 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.965225 | Mean dependent var | 3326.211 |
| Adjusted R-squared | 0.958270 | S.D. dependent var | 218.1225 |
| S.E. of regression | 44.55809 | Akaike info criterion | 10.61613 |
| Sum squared resid | 29781.36 | Schwarz criterion | 10.81496 |
| Log likelihood | -96.85322 | Durbin-Watson stat | 1.778970 |

Побудова корелограми зображено на рисунку 2.9

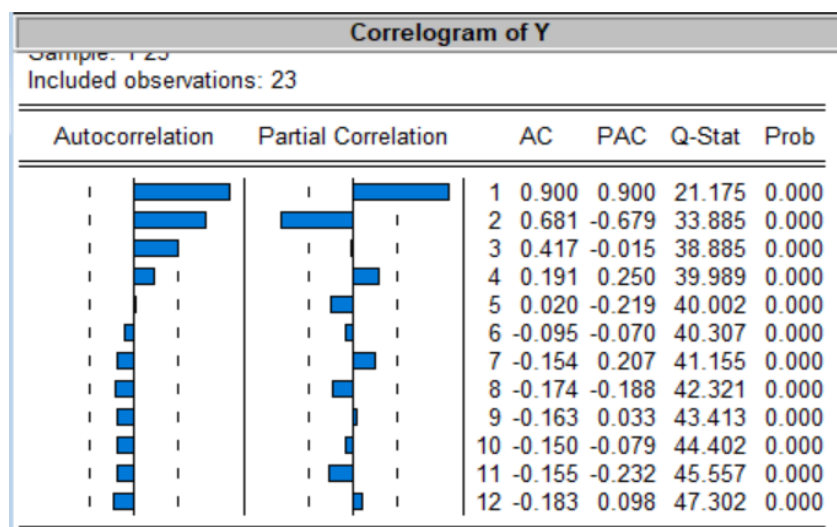


Рис. 2.9 - корелограма підключення інтернету В2С

Побудова прогнозування зображено на рисунку 2.10

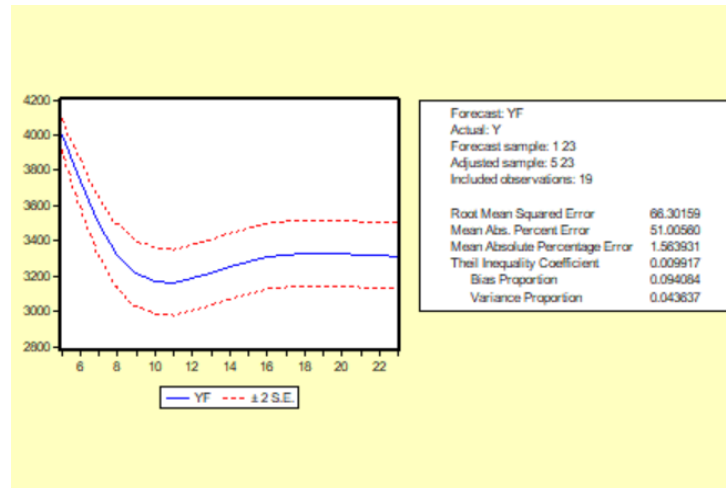


Рис. 2.10 - прогнозування підключення інтернету В2С

4. Відключення інтернету в сегменті В2С

Dependent Variable: Y

Method: Least Squares

Sample(adjusted): 5 23

Included observations: 19 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4)$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|-----------|-------------|------------|--------------------|----------|
| C(1) | 128.0959 | 96.51513 | 1.327211 | 0.2043 |
| C(2) | 1.680520 | 0.263818 | 6.369997 | 0.0000 |
| C(3) | -0.694085 | 0.247041 | -2.809591 | 0.0132 |
| C(4) | -0.031735 | 0.042303 | -0.750198 | 0.4647 |
| R-squared | 0.986486 | | Mean dependent var | 2598.500 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| Adjusted R-squared | 0.983784 | S.D. dependent var | 963.4115 |
| S.E. of regression | 122.6845 | Akaike info criterion | 12.64177 |
| Sum squared resid | 225772.4 | Schwarz criterion | 12.84060 |
| Log likelihood | -116.0968 | Durbin-Watson stat | 1.867643 |

Побудова корелограми зображено на рисунку 2.11

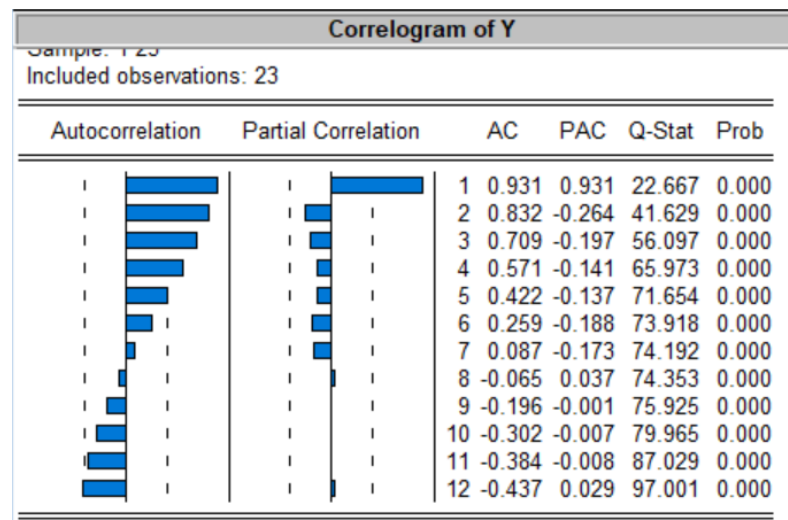


Рис. 2.11 - корелограма відключення інтернету В2С

Побудова прогнозування зображено на рисунку 2.12

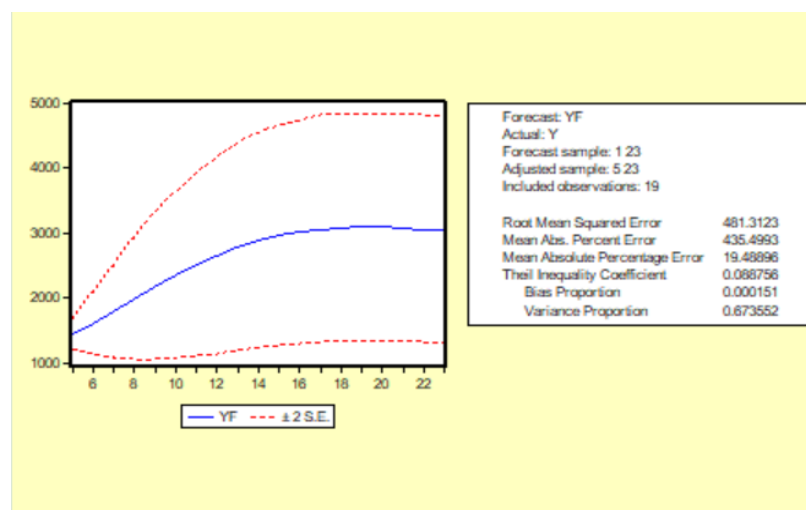


Рис. 2.12 - прогнозування відключення інтернету В2С

5. Підключення фіксованого зв'язку в сегменті B2B

Dependent Variable: Y

Method: Least Squares

Sample: 1 23

Included observations: 23

White Heteroskedasticity-Consistent Standard Errors & Covariance

$$Y = C(4) + C(5) * Y$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|--------------------|-------------|-------|
| C(4) | 0.000000 | 0.000000 | NA | NA |
| C(5) | 1.000000 | 0.000000 | NA | NA |
| R-squared | 1.000000 | Mean dependent var | 1020.348 | |
| Adjusted R-squared | 1.000000 | S.D. dependent var | 322.9842 | |
| S.E. of regression | 0.000000 | Sum squared resid | 0.000000 | |

Побудова корелограми зображено на рисунку 2.13

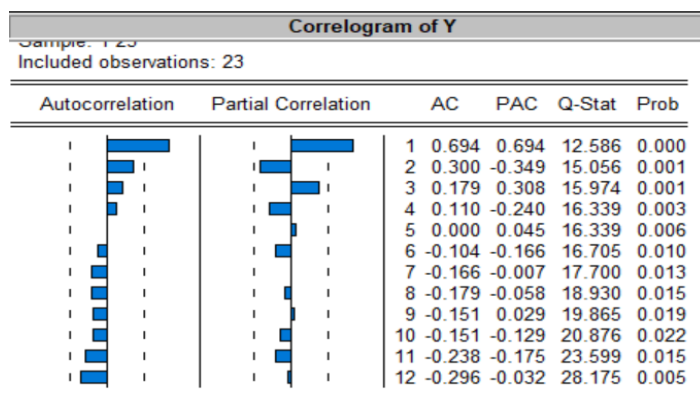


Рис. 2.13 - корелограма підключення фіксованого зв'язку B2B

Побудова прогнозування зображено на рисунку 2.14

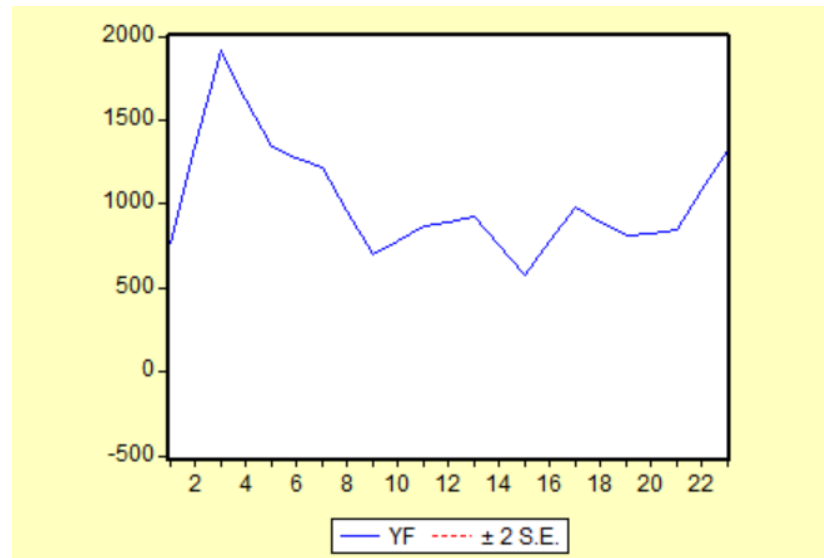


Рис. 2.14 - прогнозування підключення фіксованого зв'язку B2B

6. Відключення фіксованого зв'язку в сегменті B2B

Dependent Variable: Y

Method: ML - ARCH

Sample(adjusted): 2 23

Included observations: 22 after adjusting endpoints

Convergence achieved after 18 iterations

Bollerslev-Wooldrige robust standard errors & covariance

$$Y = C(1) + C(2) * Y(-1)$$

| | Coefficient | Std. Error | z-Statistic | Prob. |
|------|-------------|------------|-------------|--------|
| C(1) | 682.9518 | 133.5654 | 5.113239 | 0.0000 |
| C(2) | 0.776157 | 0.033024 | 23.50284 | 0.0000 |
| C(3) | 53062.71 | 15130.99 | 3.506889 | 0.0005 |
| C(4) | 0.526438 | 0.456009 | 1.154448 | 0.2483 |

| | | | | |
|---------------------|-----------|-----------|-----------------------|----------|
| C(5) | -0.499498 | 0.640878 | -0.779396 | 0.4357 |
| C(6) | -0.131427 | 0.126985 | -1.034981 | 0.3007 |
| R-squared | | 0.610980 | Mean dependent var | 3348.909 |
| Adjusted R-squared | | 0.489411 | S.D. dependent var | 481.3035 |
| S.E. of regression. | | 343.9180 | Akaike info criterion | 14.08646 |
| Sum squared resid. | | 1892473 | Schwarz criterion | 14.38401 |
| Log likelihood | | -148.9510 | Durbin-Watson stat | 0.806414 |

Побудова корелограми зображено на рисунку 2.15

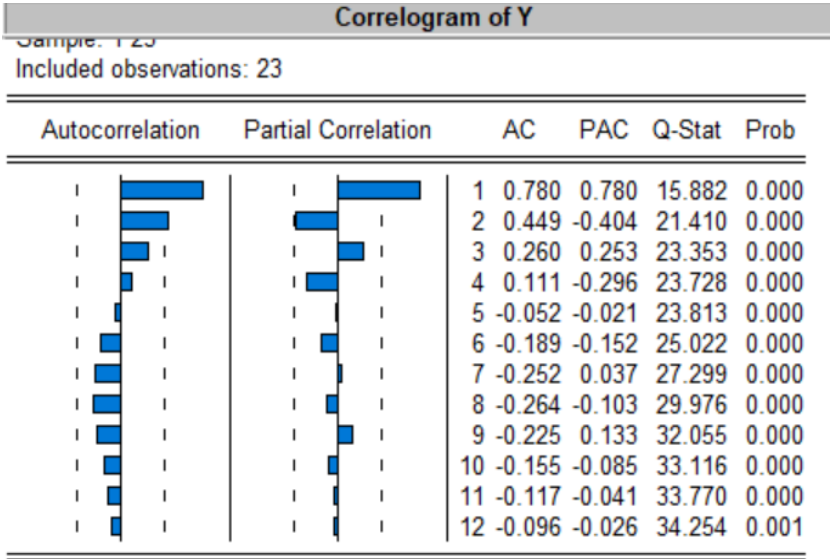


Рис. 2.15 - корелограма відключення фіксованого зв'язку В2В

Побудова прогнозування зображено на рисунку 2.16

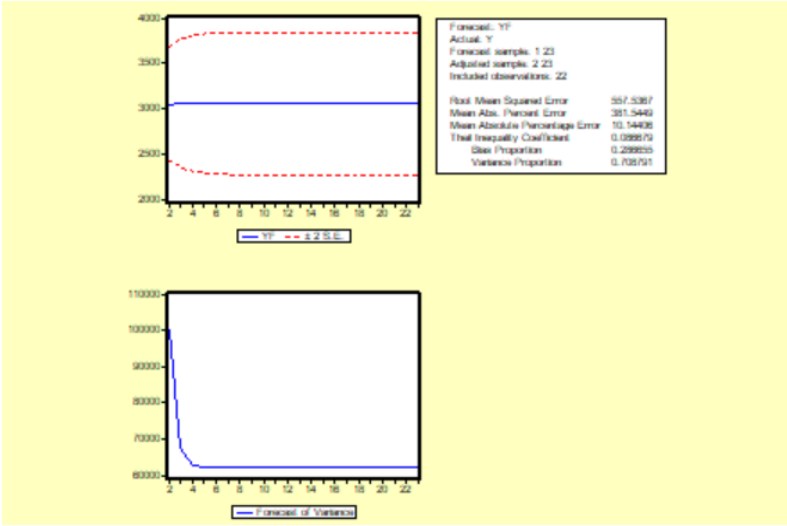


Рис. 2.16 - прогнозування відключення фіксованого зв’язку B2B

7. Підключення інтернету в сегменті B2B

Dependent Variable: Y
Method: Least Squares
Sample(adjusted): 6 23
Included observations: 18 after adjusting endpoints
White Heteroskedasticity-Consistent Standard Errors & Covariance

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4) + C(5) * Y(-5)$$

| | Coefficient | Std. Error | t-Statistic | Prob. |
|------|-------------|------------|-------------|--------|
| C(1) | 139.1046 | 80.40783 | 1.729989 | 0.1073 |
| C(2) | 1.017208 | 0.255942 | 3.974366 | 0.0016 |
| C(3) | -0.517454 | 0.215980 | -2.395844 | 0.0323 |
| C(4) | 0.031485 | 0.179780 | 0.175130 | 0.8637 |
| C(5) | 0.121605 | 0.123882 | 0.981617 | 0.3442 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.624386 | Mean dependent var | 400.6667 |
| Adjusted R-squared | 0.508812 | S.D. dependent var | 9.001634 |
| S.E. of regression | 6.308779 | Akaike info criterion | 6.751894 |
| Sum squared resid | 517.4090 | Schwarz criterion | 6.999220 |
| Log likelihood | -55.76705 | Durbin-Watson stat | 1.412139 |
| S.E. of regression | 44.55809 | Akaike info criterion | 10.61613 |
| Sum squared resid | 29781.36 | Schwarz criterion | 10.81496 |
| Log likelihood | -96.85322 | Durbin-Watson stat | 1.778970 |

Побудова корелограми зображено на рисунку 2.17

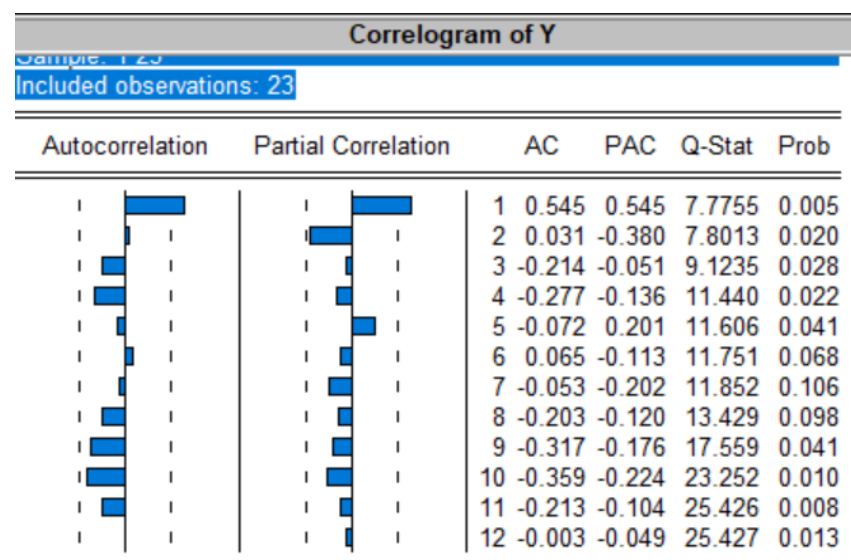


Рис. 2.17 - корелограма підключення інтернету В2В

8. Відключення інтернету в сегменті В2В

Dependent Variable: Y

Method: ML - ARCH

Sample(adjusted): 9 23

Included observations: 15 after adjusting endpoints

Convergence not achieved after 100 iterations

Bollerslev-Wooldrige robust standard errors & covariance

$$Y = C(1) + C(2) * Y(-1) + C(3) * Y(-2) + C(4) * Y(-4) + C(5) * Y(-8)$$

| | Coefficient | Std. Error | z-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C(1) | 257.6476 | 30.63796 | 8.409424 | 0.0000 |
| C(2) | 0.167516 | 0.048553 | 3.450125 | 0.0006 |
| C(3) | -0.445760 | 0.071604 | -6.225371 | 0.0000 |
| C(4) | 0.124960 | 0.052285 | 2.389962 | 0.0169 |
| C(5) | -0.118002 | 0.038723 | -3.047365 | 0.0023 |
| C(6) | 10.56672 | 4.951379 | 2.134097 | 0.0328 |
| C(7) | 1.433400 | 0.515947 | 2.778191 | 0.0055 |
| C(8) | -0.032386 | 0.029337 | -1.103936 | 0.2696 |
| R-squared | 0.066876 | Mean dependent var. | 203.0667 | |
| Adjusted R-squared | -0.866248 | S.D. dependent var | 23.56895 | |
| S.E. of regression | 32.19770 | Akaike info criterion | 8.223335 | |
| Sum squared resid | 7256.843 | Schwarz criterion | 8.600962 | |
| Log likelihood | -53.67501 | Durbin-Watson stat | 0.851305 | |

Побудова корелограми зображено на рисунку 2.18

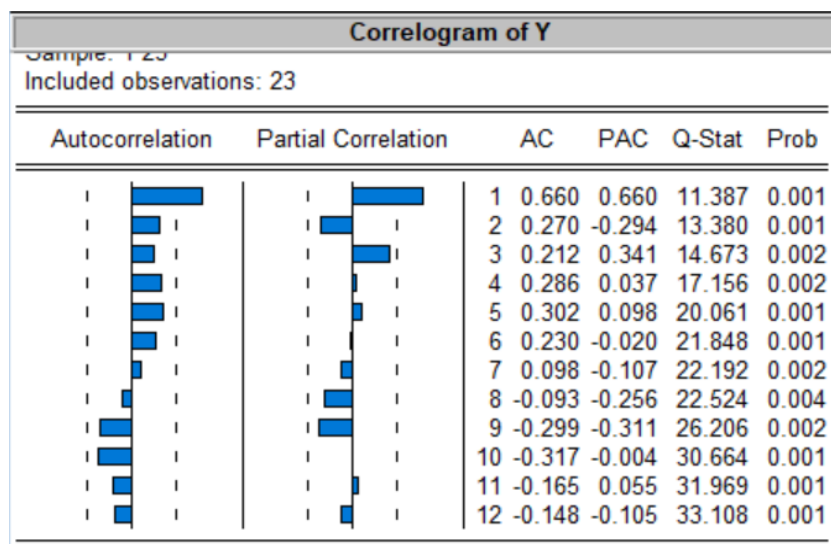


Рис. 18 - корелограма відключення інтернету В2В

Побудова прогнозування зображено на рисунку 2.19

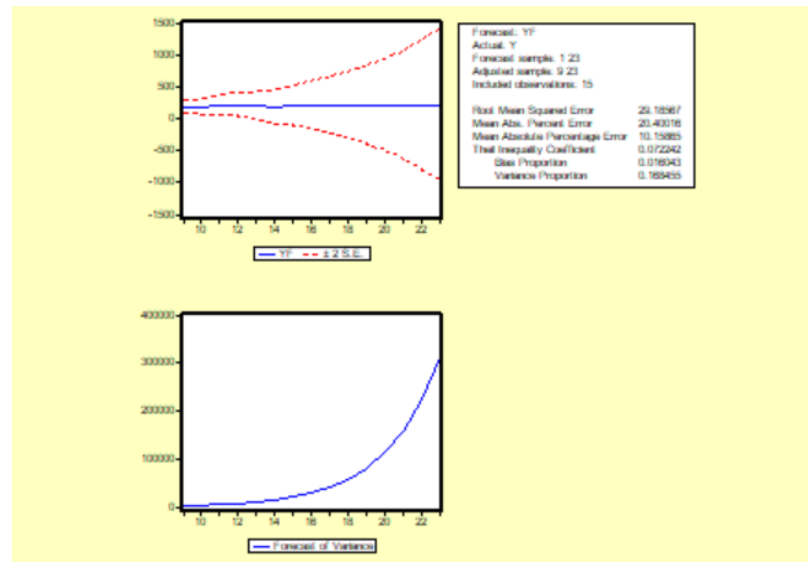


Рис. 2.19 - прогнозування відключення інтернету B2B

Висновки до розділу 2

Цей розділ присвячено аналізу методам моделювання і прогнозування процесів на телекомунікаційному ринку України. Було детально розглянуто існуючі методи моделювання та прогнозування. Також було розглянуто існуюче програмне забезпечення для прогнозування і за допомогою нього було спрогнозовані процеси і підготовлені матеріали для аналізу з моїм програмним забезпеченням, яке буде представлено у наступному розділі.

РОЗДІЛ 3 РОЗРОБКА ПРОГРАМНОГО ПРОДУКТУ ТА ЙОГО ЗАСТОСУВАННЯ

3.1 Розробка програмного продукту

У процесі виконання дипломної роботи було розроблено і практично реалізовано програмний продукт, який призначений для аналізу процесів в телекомунікаціях. А саме: підключення та відключення фіксованого зв'язку в системах B2B та B2C, та підключення та відключення дротового інтернету ADSL в системах B2C та B2B. Завдяки програмному продукту ми можемо спрогнозувати на основі даних, що може відбуватися в цих процесах у близькому майбутньому.

Головною метою створення даної програми було проілюструвати роботу методу моделювання процесів, застосування на реальних даних, та отримання відповідного прогнозу стосовно подальшого розвитку подій. Структуру програми можна представити за допомогою схеми (рис.3.1).

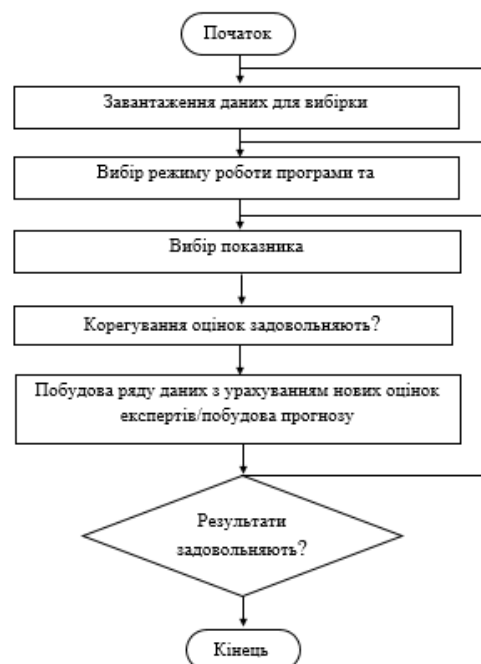


Рисунок 3.1 - Схема програми

Система підтримки прийняття рішень (СППР) — інтерактивна комп'ютерна система, що використовується з метою допомогти у прийнятті рішень в бізнес-діяльності. Такі системи дозволяють одержати важливу інформацію з вихідних даних, забезпечують проведення аналізу та надають відповідну модель з метою забезпечення вирішення поставленої задачі.

СППР, що була реалізована у даній роботі, з точки зору архітектури є простою. Після її побудови, вона може бути використана у малих компаніях з низьким рівнем інформаційних технологій.

Архітектура СППР включає в себе наступне:

- Інтерпретатор;
- Система обробки даних та генерування результатів;
- База даних і знань;
- Система прийняття рішень

3.2 Вибір інструментальної платформи для реалізації програмного продукту, інструкція по експлуатації програми

В ході виконання роботи було розглянуто декілька мов програмування для реалізації програмної частини роботи: C++, C#, Python, Matlab. В кінцевому результаті перевагу було віддано на користь Python.

Python – зручна у використанні та гнучка мова програмування, яка також дозволяє легко створювати графічний інтерфейс за допомогою різних бібліотек. Python має багато вбудованих математичних функцій та процедур, має безліч бібліотек для роботи з графіками. Великою перевагою стало те, що Python дозволяє легко оперувати матрицями та великими масивами даних без необхідності завчасного виділення під них пам'яті. А також великою

перевагою є те, що цей код можна використовувати на різних операційних системах.

3.3 Інструкція по експлуатації

Програмний продукт виконаний в двох вікнах адже кожна програма будує різні моделі. Розглянемо першу програму та інструкцію для неї. Перша програма – побудова Авторегресії та Автокореляції.

Ми маємо код без інтерфейсу, але це не проблема, адже код можна запустити в Jupiter Notebook на будь-якій операційній системі. У цьому є перевага, адже для використання цього програмного продукту потрібно дві речі. А саме:

1. Перемістити код у теку, де є файл з підготовленими даними.
2. У самому коді просто замінити назву тестового файлу, на той, що Вам потрібен. У рисунку 3.2 підсвічено, де саме потрібно вносити назву потрібного для моделювання файлу.

Приклад:

```
from pandas import read_csv
from matplotlib import pyplot
from pandas import DataFrame
from pandas import concat
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.ar_model import AutoReg
from math import sqrt

series = read_csv('B2BTELOFF.csv', header=0, index_col=0, sep=';')
print(series.head())
series.plot()
pyplot.show()
```

Рисунок 3.2 - Програмний код

Крок 1. Необхідно обрати файл для прогнозування. Файл повинен бути в розширенні .csv.

Крок 2. Дані завантажуються автоматично. Код нижче завантажить набір даних у вигляді серії Pandas. Після запуску побудується лінійний графік набору даних. Нижче у рисунках 3.3 – 3.10 зображені приклади наших даних.

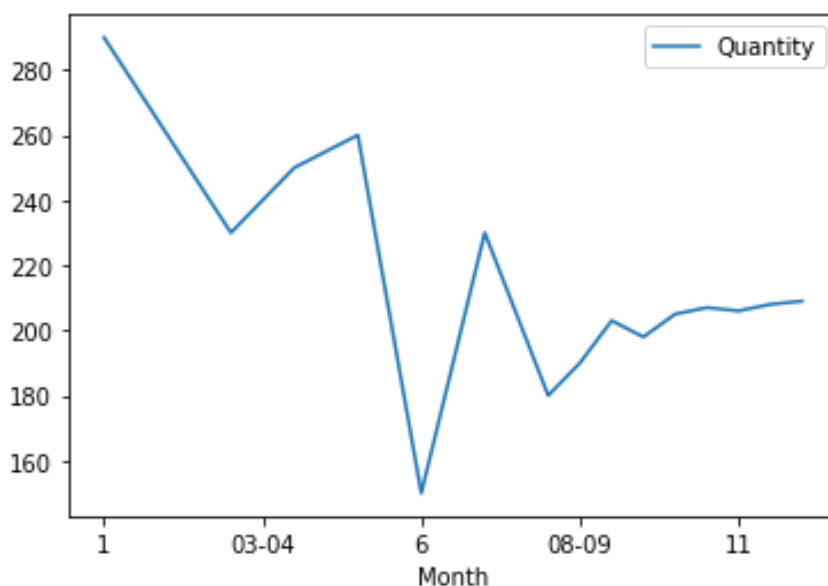


Рисунок 3.3 - Лінійний графік, побудований на даних системи В2С (підключення фіксованого зв'язку)

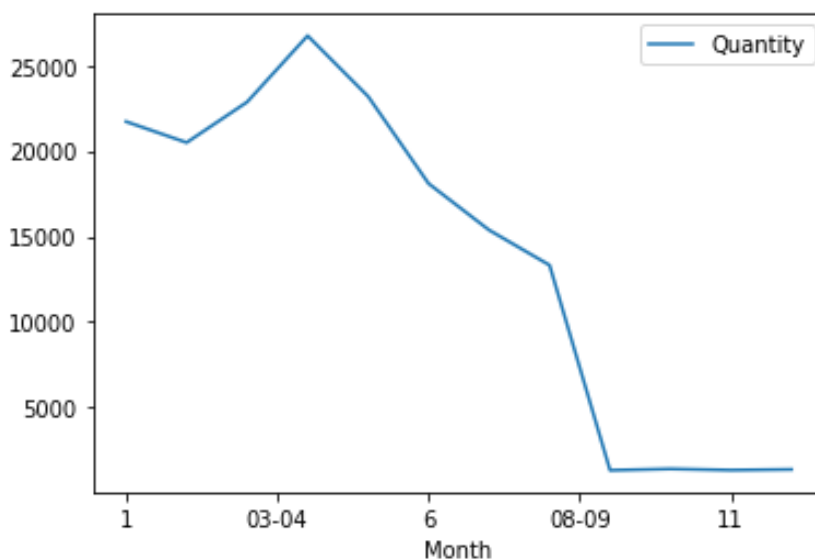


Рисунок 3.4 - Лінійний графік, побудований на даних системи B2C
(відключення фіксованого зв'язку)

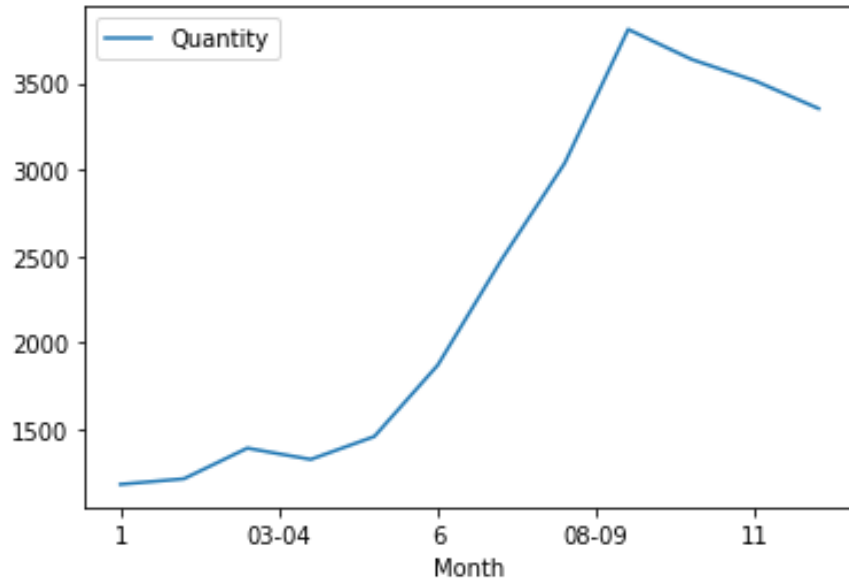


Рисунок 3.5 - Лінійний графік, побудований на даних системи B2C
(підключення інтернету ADSL)

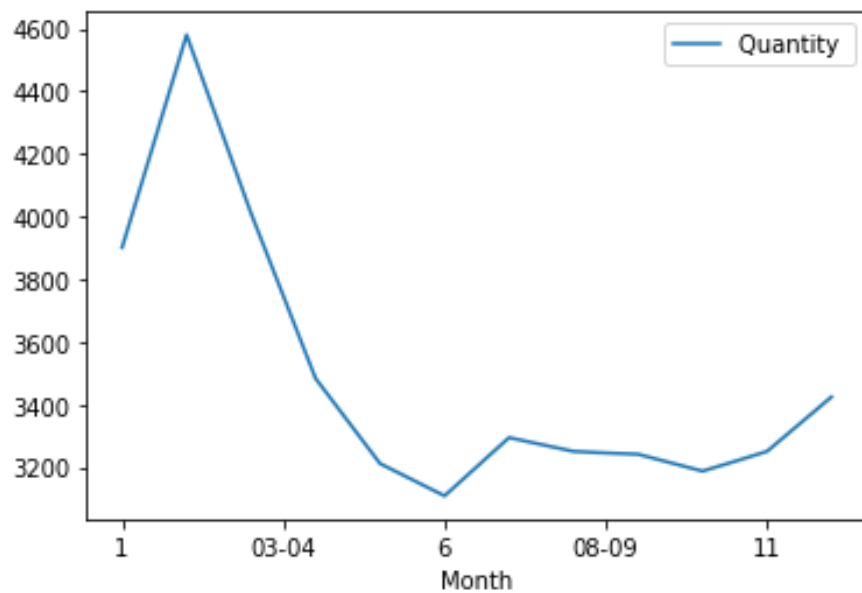


Рисунок 3.6 - Лінійний графік, побудований на даних системи B2C
(відключення інтернету ADSL)

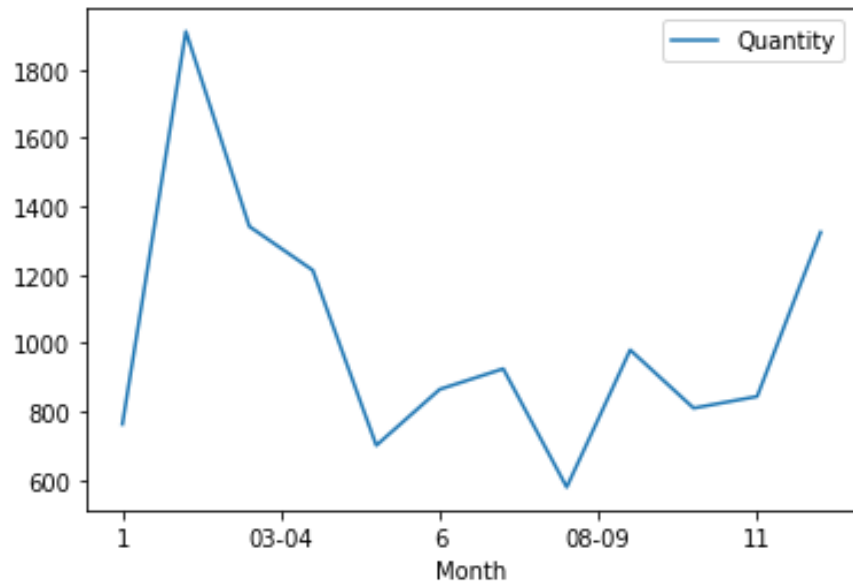


Рисунок 3.7 - Лінійний графік, побудований на даних системи В2В
(підключення фіксованого зв'язку)

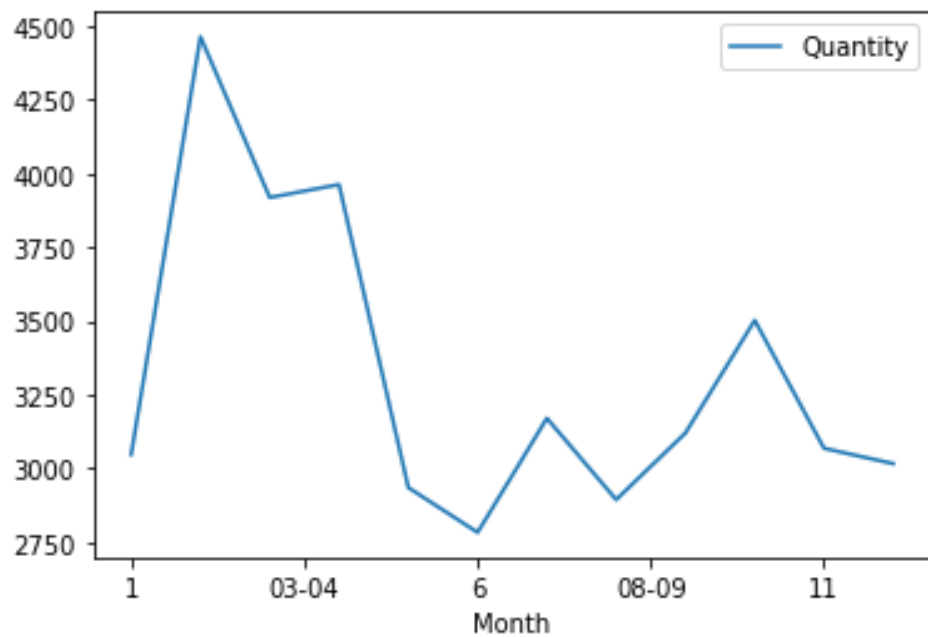


Рисунок 3.8 - Лінійний графік, побудований на даних системи В2В
(відключення фіксованого зв'язку)

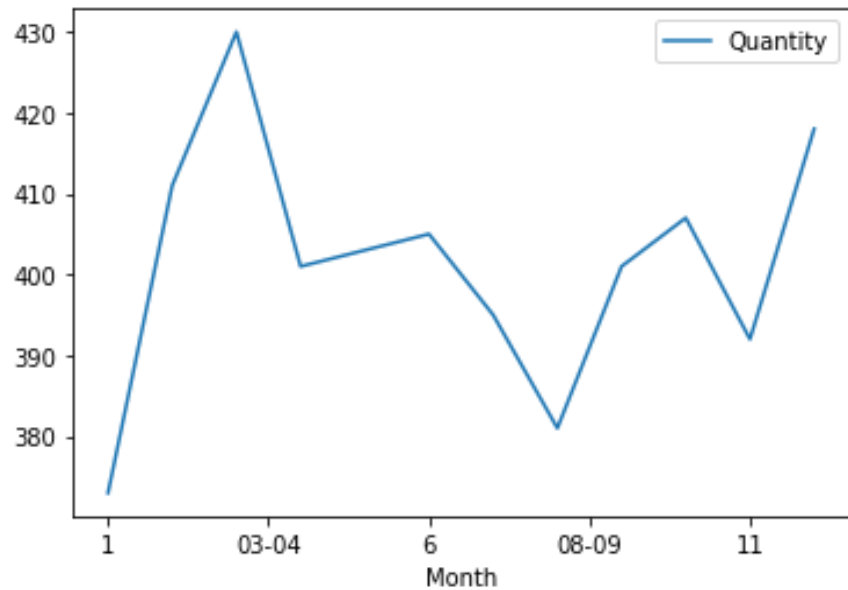


Рисунок 3.9 - Лінійний графік, побудований на даних системи B2B (підключення інтернету ADSL)

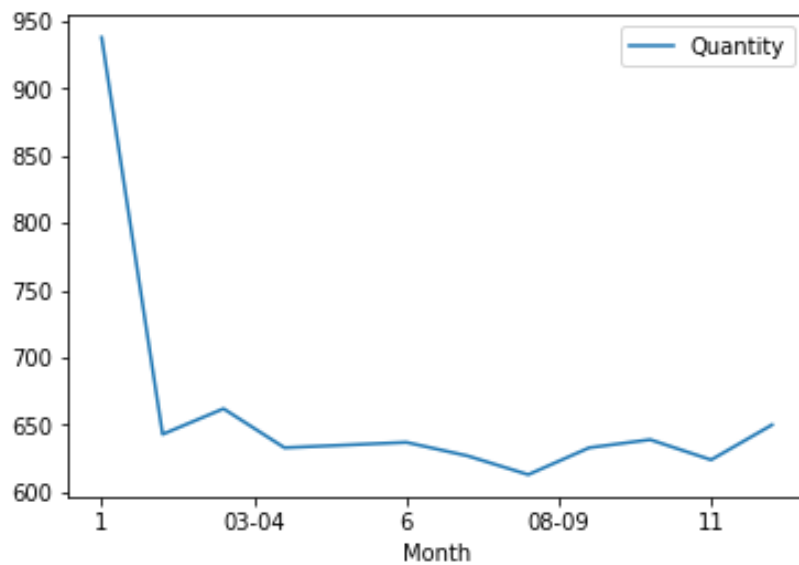


Рисунок 3.10 - Лінійний графік, побудований на даних системи B2B (відключення інтернету ADSL)

Крок 3. Побудова автокореляції. Для цього потрібно натиснути трикутник, після чого відобразиться графік даних за один місяць по осі X, і за попередній місяць по осі Y. В залежності від кількості даних, ми можемо

побачити спостереження вздовж діагональної лінії графіку, що зображені на рисунку 3.11.

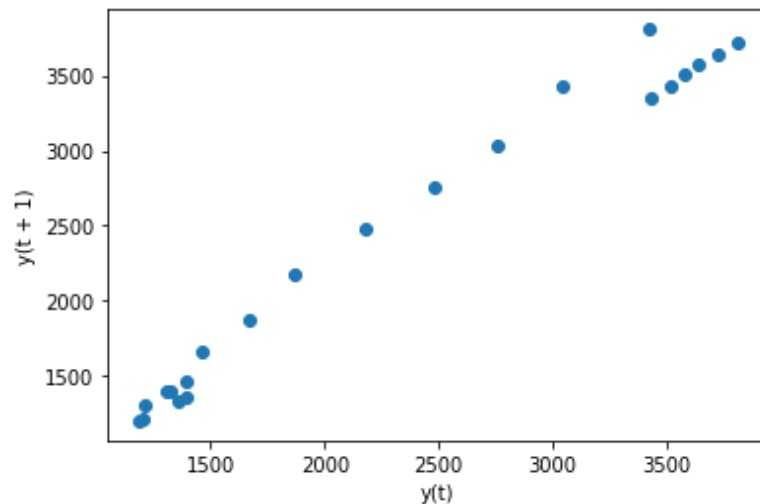


Рисунок 3.11 - Графік даних, побудований на даних системи В2С (підключення інтернету ADSL)

Крок 4. Натискаємо на трикутник на наступному блоці. Бачимо розрахунок кореляції між спостереженням та змінною «відставання». Для цього використовується тест Пірсона. Ви побачите число (від -1 (негативно корельовані) до 1 (позитивно корельовані)). Будується матриця кореляції, приклад якої можна побачити на рисунку 3.12.

| | t-1 | t+1 |
|-----|----------|----------|
| t-1 | 1.000000 | 0.986874 |
| t+1 | 0.986874 | 1.000000 |

Рисунок 3.12 - Матриця кореляції, що побудована на даних системи В2С (підключення інтернету ADSL)

Крок 5. Натискаємо на трикутник на наступному блоці, який буде графік, що показує число «запізньень» уздовж осі X і значення коефіцієнта кореляції між -1 та 1 на осі Y. Графік, що зображений на рисунку 3.13 також

включає суцільні і пунктирні лінії, які вказують 95% і 99% довірчий інтервал для значень кореляції. Значення кореляції над цими лініями є більш значущими, ніж значення під лінією, забезпечуючи поріг або обмеження для вибору найбільш значних значень «запізнювання».

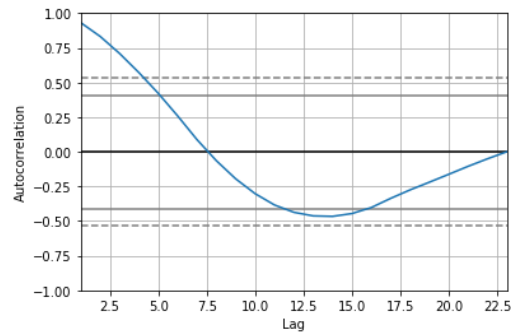


Рисунок 3.13 - Графік, який відображає число «запізнень», що побудований на даних системи В2С (підключення інтернету ADSL)

Крок 6. Натискаємо на трикутник на наступному блоці. Після запуску виводиться середня квадратична помилка (MSE), що зображена на рисунку 3.14.

Test MSE: 5908.750

Рисунок 3.14 - Відображення середньої квадратичної помилки

Крок 7. Побудова Авторгресії. При запуску виводиться графік, що зображений на рисунку 3.15.

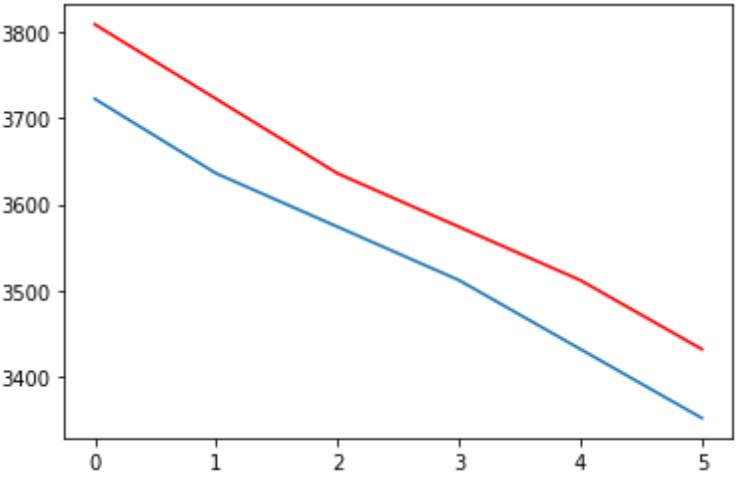


Рисунок 3.15 - Графік, який відображає очікуваних (синіх) та прогнозованих значень (червоних), що побудований на даних системи В2С (підключення інтернету ADSL)

Таблиця 3.1 - Структура вхідних даних

| Місяць | Кількість |
|--------|-----------|
| 1 | 373 |
| 01-02 | 392 |
| 2 | 411 |
| 02-03 | 420.5 |
| 3 | 430 |
| 03-04 | 415.5 |
| 4 | 401 |
| 04-05 | 402 |
| 5 | 403 |
| 05-06 | 404 |
| 6 | 405 |
| 06-07 | 400 |

| | |
|-------|-------|
| 7 | 395 |
| 07-08 | 388 |
| 8 | 381 |
| 08-09 | 391 |
| 9 | 401 |
| 09-10 | 404 |
| 10 | 407 |
| 10-11 | 399.5 |
| 11 | 392 |
| 11-12 | 405 |
| 12 | 418 |

Розглянемо другу програму та інструкцію для неї. Друга програма – побудова ARIMA-моделі та ковзаючий прогноз ARIMA.

Крок 1. Необхідно обрати файл для прогнозування. Файл повинен бути в розширенні .csv.

Крок 2. У самому коді просто замінити назву тестового файлу, на той, що Вам потрібен.

Крок 3. Натиснути на трикутник, після чого дані будуються у вигляді часових рядів з місяцем уздовж осі x та показниками на осі y, як зображено на рисунку 3.16.

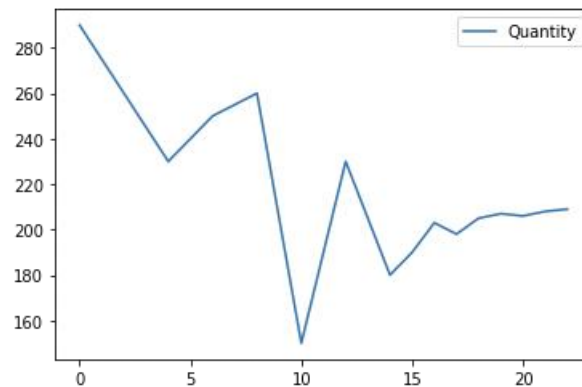


Рисунок 3.16 - Побудова тренду, на основі даних системи В2С (підключення фіксованого зв'язку)

Крок 4. Натиснути на трикутник, після чого будується графік автокореляції, який зображено на рисунку 3.17.

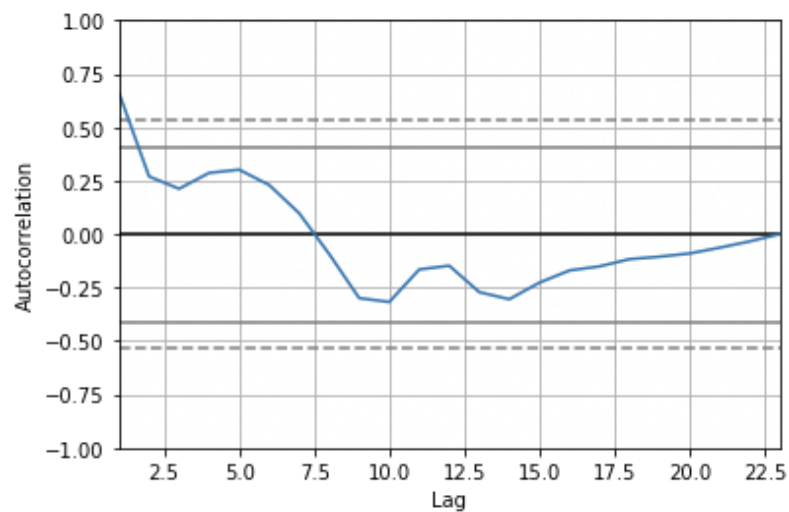


Рисунок 3.17 - Побудова автокореляції, на основі даних системи В2С (підключення фіксованого зв'язку)

Крок 4. Натиснути на трикутник, після чого будується модель ARIMA (5,1,0) (яка зображена на рисунку 3.18). Встановлюємо значення відставання в 5 для авторегресії, використовуємо різницю порядку 1, щоб зробити часовий ряд нерухомим, і використовуємо ковзну середню модель 0.

| SARIMAX Results | | | | | | |
|-------------------------|------------------|-------------------|-------------------|-------|--------|---------|
| ===== | | | | | | |
| Dep. Variable: | Quantity | No. Observations: | 23 | | | |
| Model: | ARIMA(5, 1, 0) | Log Likelihood | -90.824 | | | |
| Date: | Tue, 02 Jun 2020 | AIC | 193.647 | | | |
| Time: | 01:33:47 | BIC | 200.193 | | | |
| Sample: | 0 | HQIC | 195.189 | | | |
| | - 23 | | | | | |
| Covariance Type: | opg | | | | | |
| ===== | | | | | | |
| | coef | std err | z | P> z | [0.025 | 0.975] |
| ----- | | | | | | |
| ar.L1 | 0.6473 | 0.244 | 2.648 | 0.008 | 0.168 | 1.126 |
| ar.L2 | -1.1166 | 0.293 | -3.811 | 0.000 | -1.691 | -0.542 |
| ar.L3 | 0.6580 | 0.454 | 1.448 | 0.148 | -0.233 | 1.549 |
| ar.L4 | -0.5309 | 0.373 | -1.423 | 0.155 | -1.262 | 0.200 |
| ar.L5 | 0.2593 | 0.323 | 0.802 | 0.423 | -0.375 | 0.893 |
| sigma2 | 202.9423 | 61.523 | 3.299 | 0.001 | 82.359 | 323.525 |
| ===== | | | | | | |
| Ljung-Box (Q): | | 4.31 | Jarque-Bera (JB): | | 18.12 | |
| Prob(Q): | | 1.00 | Prob(JB): | | 0.00 | |
| Heteroskedasticity (H): | | 0.51 | Skew: | | -1.44 | |
| Prob(H) (two-sided): | | 0.39 | Kurtosis: | | 6.38 | |
| ===== | | | | | | |

Рисунок 3.18 – модель ARIMA, що побудована на основі даних системи В2С (підключення фіксованого зв'язку)

Крок 5. Натиснути на трикутник, після чого ми отримуємо лінійний графік залишкових помилок, що дозволяє припустити, що все ще може бути якась інформація про тренд, не зафіксована моделлю, як зображено на рисунку 3.19.

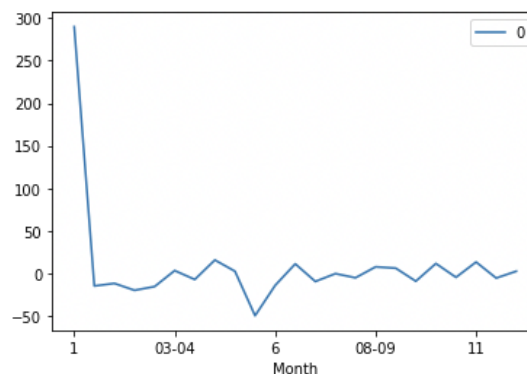


Рисунок 3.19 – графік залишкових помилок що побудовані на основі даних системи В2С (підключення фіксованого зв'язку)

Крок 6. Натиснути на трикутник, після чого ми отримуємо графік щільності залишкових значень помилок, припускаючи, що помилки є гауссовими, але не можуть бути зосереджені на нулі. Цей графік зображено на рисунку 3.20.

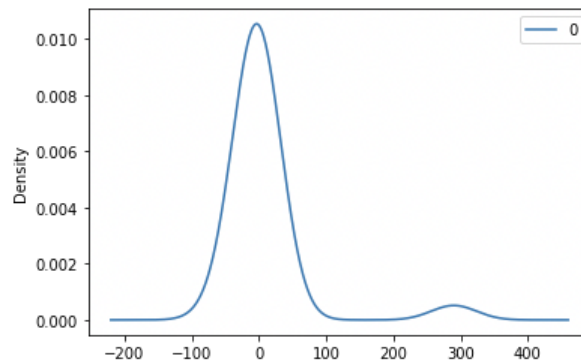


Рисунок 3.20 – графік щільності залишкових значень помилок що побудовані на основі даних системи В2С (підключення фіксованого зв’язку)

Крок 7. Натиснути на трикутник, після чого ми ми можемо обчислити підсумковий середньоквадратичний показник помилок (MSE) для прогнозів, забезпечуючи точку порівняння для інших конфігурацій ARIMA. Створюється лінійний графік, який показує очікувані значення (синім кольором) в порівнянні з прогнозами змінного прогнозу (червоним кольором). На рисунку 3.21 ми можемо бачити, що значення показують деяку тенденцію і знаходяться в правильному масштабі.

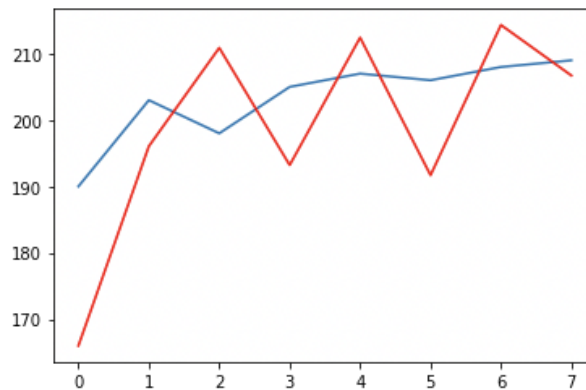


Рисунок 3.21 – графік який показує очікувані значення (синім кольором) в порівнянні з прогнозами змінного прогнозу (червоним кольором), що побудований на основі даних системи В2С (підключення фіксованого зв'язку)

Програми дозволяють будувати наступні моделі на вибір:

- AR2 (n-ного порядку);
- ARIMA (авторегресія n-ного порядку з інтегровним ковзним середнім k-ого порядку);

Висновки до розділу 3

Авторегресія - модель часових рядів, яка використовує спостереження з попередніх етапів часу як вхід до рівняння регресії для прогнозування значення на наступному кроці часу.

Це дуже проста ідея, яка може призвести до точних прогнозів щодо ряду проблем часових рядів.

Модель авторегресії робить припущення, що спостереження на попередніх етапах часу корисні для прогнозування значення на наступному кроці часу.

Цей взаємозв'язок між змінними називається кореляцією.

Якщо обидві змінні змінюються в одному напрямку (наприклад, разом йти разом або вниз), це називається позитивною кореляцією. Якщо змінні рухаються в протилежних напрямках у міру зміни значень (наприклад, одна йде вгору, а одна знижується), то це називається негативною кореляцією.

Ми можемо використовувати статистичні заходи для обчислення кореляції між вихідною змінною та значеннями на попередніх етапах часу при різних відставаннях. Чим сильніша кореляція між вихідною змінною та конкретною відсталою змінною, тим більше ваги, яку модель авторегресії може надати цій змінній при моделюванні.

Знову ж таки, оскільки кореляція обчислюється між змінною та самою на попередніх етапах часу, вона називається автокореляцією. Його також називають послідовною кореляцією через послідовність структури даних часових рядів.

Кореляційна статистика також може допомогти вибрати, які змінні відставання будуть корисні в моделі, а які ні.

Цікаво, що якщо всі змінні затримки показують низьку кореляцію або відсутність кореляції з вихідною змінною, то це говорить про те, що проблема часових рядів може бути непередбачуваною. Це може бути дуже корисно для початку роботи з новим набором даних.

Модель авторегресії - модель лінійної регресії, яка використовує відсталі змінні як вхідні змінні.

Ми обчислили модель лінійної регресії вручну, використовуючи клас `LinearRegression` в `scikit-learn` і вручну вказали вхідні змінні вхідних відстань.

Крім того, бібліотека статистичних моделей забезпечує модель авторегресії, де потрібно вказати відповідне значення відставання та тренувати лінійну регресійну модель. Він надається в класі `AutoReg`.

Ми можемо використовувати цю модель, спочатку створивши модель `AutoReg ()`, а потім викликаючи `fit ()`, щоб навчити її на нашому наборі даних. Це повертає об'єкт `AutoRegResults`.

Як тільки підходить, ми можемо використовувати модель для прогнозування, викликаючи функцію `predict ()` для низки спостережень у майбутньому. Це створює 1 7-денний прогноз, який відрізняється від прикладу стійкості вище.

РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ

4.1 Постановка задачі

У роботі застосовується метод ФВА для проведення техніко-економічний аналізу розробки.

Відповідно цьому варто обирати і систему показників якості програмного продукту.

Технічні вимоги до продукту наступні:

- програмний продукт повинен функціонувати на персональних комп'ютерах із стандартним набором компонент;
- забезпечувати високу швидкість обробки великих об'ємів даних у реальному часі;
- забезпечувати зручність і простоту взаємодії з користувачем або з розробником програмного забезпечення у випадку використання його як модуля;
- передбачати мінімальні витрати на впровадження програмного продукту.

4.1.1 Обґрунтування функцій програмного продукту

Головна функція F_0 – розробка програмного продукту, який аналізує процес за вхідними даними та будує його модель для подальшого прогнозування. Виходячи з конкретної мети, можна виділити наступні основні функції ПП:

F_1 – вибір мови програмування;

F_2 – середовище розробки.

Кожна з основних функцій може мати декілька варіантів реалізації.

Функція F_1 :

- а) мова програмування MatLab;
- б) мова програмування Python;
- в) мова програмування C#.

Функція F_2 :

- а) середовище розробки Jupyter;
- б) середовище розробки PyCharm;
- в) середовище розробки Visual Code.

4.1.2 Варіанти реалізації основних функцій

Варіанти реалізації основних функцій наведені у морфологічній карті системи (рис. 4.1). На основі цієї карти побудовано позитивно-негативну матрицю варіантів основних функцій (таблиця 4.1).

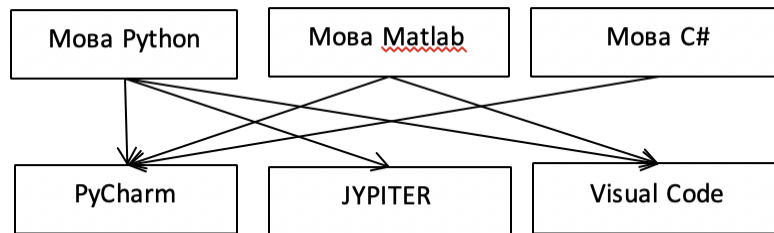


Рисунок 4.1 – Морфологічна карта

Морфологічна карта відображує всі можливі комбінації варіантів реалізації функцій, які складають повну множину варіантів ПП.

Таблиця 4.1 – Позитивно-негативна матриця

| Осно вні функції | Варі анти реалізації | Переваги | Недоліки |
|------------------------|----------------------------|---|---|
| <i>F1</i> | <i>A</i> | Перевірка результатів в режимі реального часу, висока швидкодія, кросплатформеність, широкий спектр функцій та ООП можливості | Висока вартість ліцензії, високі вимоги до системи |
| | <i>B</i> | Займає менше часу при написанні коду, має усі необхідні бібліотеки для цього проекту, кросплатформений | Потрібно вміти користуватися необхідними бібліотеками |
| | <i>B</i> | ООП | Не кросплатформений та займає більше часу для написання коду через відсутність вбудованого функціоналу математичних обчислень |
| <i>F2</i> | <i>A</i> | Простота створення | Обмежений функціонал |
| | <i>B</i> | Простота створення | Відсутність кросплатформеності |
| | <i>B</i> | Кросплатформенний | Складність у використанні |

На основі аналізу позитивно-негативної матриці робимо висновок, що при розробці програмного продукту деякі варіанти реалізації функцій варто

відкинути, тому, що вони не відповідають поставленим перед програмним продуктом задачам. Ці варіанти відзначені у морфологічній карті.

Функція $F1$:

Оскільки розрахунки проводяться з використанням складного матапарату, то має бути відкинтий варіант А, а також через значно простішу для розуміння документацію Python відкидаємо варіант Б.

Функція $F2$:

Оскільки, програмний продукт реалізується на мові Python, відкидаємо варіант В через високий рівень складності реалізації.

Таким чином, будемо розглядати такі варіанти реалізації ПП:

$F1a - F2a$

1. $F1a - F2b$

Для оцінювання якості розглянутих функцій обрана система параметрів, описана нижче.

4.2 Обґрунтування системи параметрів ПП

4.2.1 Опис параметрів

Для того, щоб охарактеризувати програмний продукт, будемо використовувати наступні параметри:

- $X1$ – швидкодія мови програмування;
- $X2$ – об'єм пам'яті для збереження даних;
- $X3$ – час обробки даних;
- $X4$ – потенційний об'єм програмного коду,

де $X1, X2, X3$ – параметри функції $F1$, $X4$ – параметр функції $F2$.

4.2.2 Кількісна оцінка параметрів

Гірші, середні і кращі значення параметрів вибираються на основі вимог замовника й умов, що характеризують експлуатацію ПП як показано у табл. 4.2.

Таблиця 4.2 – Основні параметри ПП

| НазваПараметра | Умовні позначення | Одиниці виміру | Значення параметра | | |
|------------------------------------|-------------------|----------------------|--------------------|----------|-------|
| | | | гірш і | серед ні | кращі |
| Швидкодія мови програмування | X1 | Оп/мс | 2000 | 11000 | 19000 |
| Об'єм пам'яті для збереження даних | X2 | Мб | 32 | 16 | 8 |
| Час обробки запитів користувача | X3 | мс | 1000 | 420 | 60 |
| Потенційний об'єм програмного коду | X4 | кількість строк коду | 1000 | 600 | 300 |

За даними таблиці 4.2 будуються графічні характеристики параметрів (рис. 4.2 – рис. 4.5).

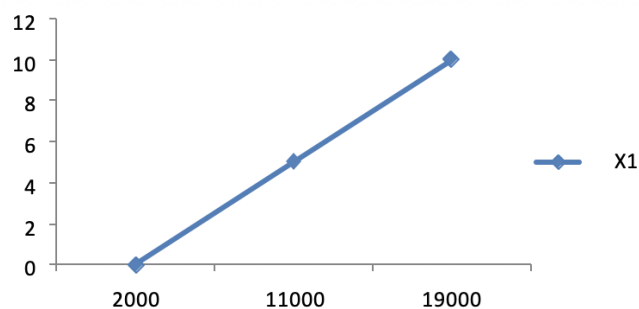


Рисунок 4.2 – X1, швидкодія мови програмування

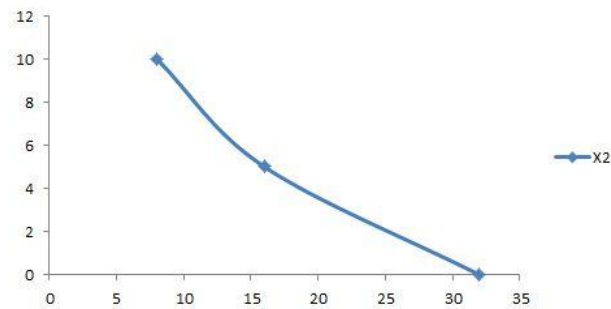


Рисунок 4.3 – X2, об'єм пам'яті для збереження даних

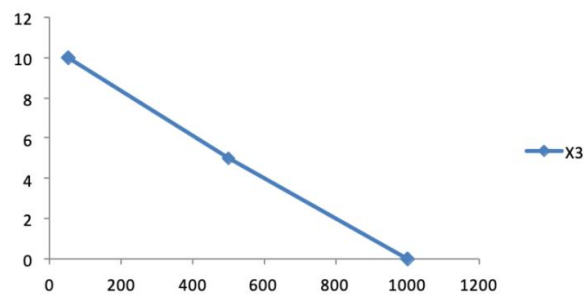


Рисунок 4.4 – X3, час виконання запитів користувача

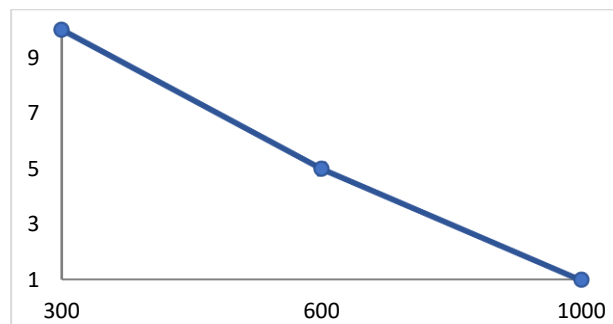


Рисунок 4.5 – X4, потенційний об'єм програмного коду

4.2.3 Аналіз експертного оцінювання параметрів

Значимість кожного параметра визначається методом попарного порівняння. Оцінку проводить експертна комісія із 7 людей. Результати експертного ранжування наведені у таблиці 4.3.

Таблиця 4.3 – Результати ранжування параметрів

| Позначення параметра | Назва параметра | Одиниці виміру | Ранг параметра за оцінкою експерта | | | | | | | Сума рангів R_i | Відхилення Δ_i | Δi^2 |
|----------------------|------------------------------------|---------------------|------------------------------------|----|----|----|----|----|----|-------------------|-----------------------|--------------|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | |
| X1 | Швидкодія мови програмування | Оп/мс | 2 | 2 | 2 | 3 | 2 | 2 | 2 | 15 | -2.5 | 6.25 |
| X2 | Об'єм пам'яті для збереження даних | Мб | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 8 | -9.5 | 90.25 |
| X3 | Час обробки запитів користувача | мс | 3 | 3 | 4 | 1 | 3 | 3 | 4 | 21 | 3.5 | 12.25 |
| X4 | Потенційний об'єм програмного коду | кількість строккоду | 4 | 4 | 3 | 4 | 4 | 4 | 3 | 26 | 8.5 | 72.25 |
| | Разом | | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 70 | 0 | 181 |

Порахуємо коефіцієнт узгодженості:

$$W = \frac{12S}{N^2(n^3 - n)} = \frac{12 \cdot 181}{7^2(4^3 - 4)} = 0,739 > W_k = 0.67$$

Ранжування можна вважати достовірним, тому що знайдений коефіцієнт узгодженості перевищує нормативний, котрий дорівнює 0.67.

Скориставшись результатами ранжирування, проведемо попарне порівняння всіх параметрів і результати занесемо у таблицю 4.4. За найвищий ранг приймається 4.

Таблиця 4.4 – Попарне порівняння параметрів

| Параметри | Експерти | | | | | | | Кінцева оцінка | Числове значення |
|-----------|----------|---|---|---|---|---|---|----------------|------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| X1 і X2 | > | > | < | > | > | > | > | > | 1,5 |
| X1 і X3 | < | < | > | < | < | < | < | < | 0,5 |
| X1 і X4 | < | < | < | > | < | < | < | < | 0,5 |
| X2 і X3 | < | < | < | < | < | < | < | < | 0,5 |
| X2 і X4 | < | < | < | < | < | < | < | < | 0,5 |
| X3 і X4 | < | < | < | > | < | < | > | < | 0,5 |

Як видно з таблиці 4.5, різниця значень коефіцієнтів вагомості не перевищує 2%, тому більшої кількості ітерацій не потрібно.

Таблиця 4.5 – Розрахунок вагомості параметрів

| Параметри x_i | Параметри x_j | | | | Перша ітер. | | Друга ітер. | |
|-----------------|-----------------|---------|---------|---------|-------------|-------------|-------------|------------|
| | X1 | X2 | X3 | X4 | b_i | K_{bi} | b_i^1 | K_{bi}^1 |
| X1 | 1. 0 | 1. 5 | 0. 5 | 0. 5 | 3.5 | 0.2187 5 | 12.2 5 | 0.177 5 |
| X2 | 0. 5 | 1. 0 | 0. 5 | 0. 5 | 2.5 | 0.1562 5 | 6.25 | 0.190 5 |
| X3 | 1. 5 | 1. 5 | 1. 0 | 0. 5 | 4.5 | 0.2812 5 | 20.2 5 | 0.293 4 |
| X4 | 1. 5 | 1. 5 | 1. 5 | 1. 0 | 5.5 | 0.3437 5 | 30.2 5 | 0.338 4 |
| Всього: | | | | | 16 | 1 | 98 | 1 |

4.3 Аналіз рівня якості варіантів реалізації функцій

Визначаємо рівень якості кожного варіанту виконання основних функцій.

Таблиця 4.6 – Розрахунок показників рівня якості варіантів реалізації основних функцій ПП

| Основні функції | Варіант реалізації функції | Абсолютне значення параметра | Бальна оцінка параметра | Коефіцієнт вагомості параметра | Коефіцієнт рівня якості |
|-----------------|----------------------------|------------------------------|-------------------------|--------------------------------|-------------------------|
| F1(X1, X2, X3) | Python | X1: 11000 | 5 | 0.1775 | 0.8875 |
| | | X2: 16 | 5 | 0.1905 | 0.9525 |
| | | X3: 500 | 4.45 | 0.2934 | 1.3056 |
| F2(X4) | VSCode | X4: 900 | 2.7 | 0.3384 | 0.9137 |
| | Jupyter | X4: 400 | 8.3 | 0.3384 | 2.8088 |

За даними з таблиці 4.6 визначаємо рівень якості кожного з варіантів:

$$K_{K1} = 0.8875 + 0.9525 + 1.3056 + 0.9137 = 4.06$$

$$K_{K2} = 0.8875 + 0.9525 + 1.3056 + 2.8088 = 5.95$$

Як видно з розрахунків, кращим є перший варіант, для якого коефіцієнт технічного рівня має найбільше значення.

4.4 Економічний аналіз варіантів розробки ПП

Для визначення вартості розробки ПП спочатку проведемо розрахунок трудомісткості.

Всі варіанти включають в себе два окремих завдання:

1. Розробка проекту програмного продукту;
2. Розробка програмної оболонки;

В свою чергу, варіант II містить завдання:

3. Реалізація алгоритму, що виводить результат з графічною ілюстрацією та розгорнутою інформацією.

Крім того, варіант I містить завдання:

4. Реалізація алгоритму, виводу тексту на екран програмного продукту

Завдання 1 за ступенем новизни відноситься до групи А, завдання 2 – до групи Б, завдання 3, 4 – до групи В. За складністю алгоритми, які використовуються в завданні 1 належать до групи 1; а в завданні 2 – до групи 3, завдання 3 – до групи 2, а завдання 4 - до групи 3.

Для першого завдання (ступінь новизни А, група складності – 1), тобто $T_p=90$ людино-днів, $K_{\Pi} = 1.7$, $K_{СК} = 1$; $K_{СТ} = 0.8$.

$$T_1 = 90 \cdot 0.8 \cdot 1.7 = 122.4 \text{ людино-днів.}$$

Для другого завдання (ступінь новизни Б, група складності – 3), тобто $T_p=19$ людино-днів, $K_{\Pi}=0.9$, $K_{СК} = 1$, $K_{СТ}=0.8$:

$$T_2 = 19 \cdot 0.8 \cdot 0.9 = 13.68 \text{ людино-днів.}$$

Для третього завдання (ступінь новизни В, група складності – 2), тобто $T_p=19$ людино-днів, $K_{II} = 1.2$; $K_{СК} = 1$; $K_{СТ} = 0.6$:

$$T_3 = 19 \cdot 1.2 \cdot 1 \cdot 0.6 = 13.68 \text{ людино-днів.}$$

Для четвертого завдання (ступінь новизни В, група складності – 3), тобто $T_p=12$ людино-днів, $K_{II} = 0.8$; $K_{СК} = 1$; $K_{СТ} = 0.6$:

$$T_4 = 12 \cdot 0.8 \cdot 1 \cdot 0.6 = 5.76 \text{ людино-днів.}$$

Складаємо трудомісткість завдання для кожного з обраних варіантів реалізації програми:

$$T_I = (122.4 + 13.68 + 13.68) \cdot 6 = 898.56 \text{ людино-годин;}$$

$$T_{II} = (122.4 + 13.68 + 5.76) \cdot 6 = 851.04 \text{ людино-годин;}$$

Більш високу трудомісткість має варіант I.

В розробці бере участь один програміст з окладом 12000 грн. Визначимо зарплату за годину за формулою:

$$C_{\text{ч}} = \frac{12000}{1 \cdot 21 \cdot 8} = 71.43 \text{ грн.}$$

Тоді, розрахуємо заробітну плату за формулою

Зарплата розробників за варіантами становить:

$$\text{I. } C_{\text{ЗП1}} = 71.43 \cdot 898.56 \cdot 1.2 = 77020.97 \text{ грн.}$$

$$\text{II. } C_{\text{ЗП2}} = 71.43 \cdot 851.04 \cdot 1.2 = 72947.75 \text{ грн.}$$

Відрахування на єдиний соціальний внесок в залежності від групи професійного ризику (II клас) становить 22%:

$$\text{I. } C_{\text{ВІД1}} = C_{\text{ЗП}} \cdot 0.22 = 77020.97 \cdot 0.22 = 16944.61 \text{ грн.}$$

$$\text{II. } C_{\text{ВІД2}} = C_{\text{ЗП}} \cdot 0.22 = 72947.75 \cdot 0.22 = 16048.5 \text{ грн.}$$

Тепер визначимо витрати на оплату однієї машино-години. (C_M)

Так як одна ЕОМ обслуговує одного програміста з окладом 12000 грн., з коефіцієнтом зайнятості 0,2 то для однієї машини отримаємо:

$$C_T = 12 \cdot M \cdot K_3 = 12 \cdot 12000 \cdot 0.2 = 28800 \text{ грн.}$$

З урахуванням додаткової заробітної плати:

$$C_{3П} = C_{Г} \cdot (1 + K_3) = 28800 \cdot (1 + 0.2) = 34560 \text{ грн.}$$

Відрахування на єдиний соціальний внесок:

$$C_{ВІД} = C_{3П} \cdot 0.22 = 34560 \cdot 0.22 = 7603.2 \text{ грн.}$$

Амортизаційні відрахування розраховуємо при амортизації 25% та вартості ЕОМ – 10000 грн.

$$C_A = K_{ТМ} \cdot K_A \cdot Ц_{ПР} = 1.15 \cdot 0.25 \cdot 10000 = 2875 \text{ грн.},$$

Витрати на ремонт та профілактику розраховуємо як:

$$C_P = K_{ТМ} \cdot Ц_{ПР} \cdot K_P = 1.15 \cdot 10000 \cdot 0.05 = 575 \text{ грн.},$$

Ефективний годинний фонд часу персонального комп'ютеру за рік розраховуємо за формулою:

$$T_{ЕФ} = (D_K - D_B - D_C - D_P) \cdot t_3 \cdot K_B = (366 - 118 - 16) \cdot 8 \cdot 0.9 = 1670.4$$

годин,

Витрати на оплату електроенергії розраховуємо за формулою:

$$C_{ЕЛ} = T_{ЕФ} \cdot N_C \cdot K_3 \cdot Ц_{ЕН} = 1670.4 \cdot 0.156 \cdot 0.9733 \cdot 1.75 = 443.73 \text{ грн.},$$

Накладні витрати розраховуємо за формулою:

$$C_H = Ц_{ПР} \cdot 0.67 = 10000 \cdot 0.67 = 6700 \text{ грн.}$$

Тоді, річні експлуатаційні витрати будуть:

$$C_{ЕКС} = C_{3П} + C_{ВІД} + C_A + C_P + C_{ЕЛ} + C_H$$

$$C_{ЕКС} = 34560 + 7603.2 + 2875 + 575 + 443.73 + 6700 = 52756.93 \text{ грн.}$$

Собівартість однієї машино-години ЕОМ дорівнюватиме:

$$C_{М-Г} = C_{ЕКС} / T_{ЕФ} = 52756.93 / 1670.4 = 31.58 \text{ грн/час.}$$

Витрати на оплату машинного часу, в залежності від обраного варіанта реалізації, складає:

$$\text{I. } C_{M1} = C_{М-Г} \cdot T_1 = 31.58 \cdot 882.12 = 27857.35 \text{ грн.};$$

$$\text{II. } C_{M2} = C_{М-Г} \cdot T_2 = 31.58 \cdot 851.04 = 26875.84 \text{ грн.};$$

Накладні витрати складають 67% від заробітної плати:

$$\text{I. } C_{H1} = 77020.97 \cdot 0.67 = 51604.05 \text{ грн.};$$

$$\text{II. } C_{H2} = 72947.75 \cdot 0.67 = 48874.9 \text{ грн.};$$

Отже, вартість розробки програмного продукту за варіантами становить:

I. $C_{ПП1} = 77020.97 + 16944.61 + 27857.35 + 51604.05 = 173426.98$ грн.;

II. $C_{ПП2} = 72947.75 + 16048.5 + 26875.84 + 48874.9 = 164746.99$ грн.;

4.5 Вибір кращого варіанта ПП техніко-економічного рівня

Розрахуємо коефіцієнт техніко-економічного рівня за формулою:

$$K_{\text{ТЕР}1} = 4.06 / 173426.98 = 2.65 \cdot 10^{-5};$$

$$K_{\text{ТЕР}2} = 5.95 / 164746.99 = 3.61 \cdot 10^{-4};$$

Як бачимо, найбільш ефективним є другий варіант реалізації програми з коефіцієнтом техніко-економічного рівня

$$K_{\text{ТЕР}2} = 3.61 \cdot 10^{-5}.$$

Висновки до 4 розділу

В даному розділі проведено повний функціонально-вартісний аналіз програмного продукту, який було розроблено в рамках дипломного проекту. Процес аналізу можна умовно розділити на дві частини.

В першій з них проведено дослідження програмного продукту з технічної точки зору: було визначено основні функції програмного продукту та сформовано множину варіантів їх реалізації; на основі обчислених значень параметрів, а також експертних оцінок їх важливості, було обчислено

коефіцієнт технічного рівня, який і дав змогу визначити оптимальну з технічної точки зору альтернативу реалізації функцій програмного продукту.

Другу частину ФВА присвячено вибору із альтернативних варіантів реалізації найбільш економічно обґрунтованого. Порівняння запропонованих варіантів реалізації в рамках даної частини виконувалось за коефіцієнтом ефективності, для обчислення якого були обчислені такі допоміжні параметри, як трудомісткість, витрати на заробітну плату, накладні витрати.

Після виконання функціонально-вартісного аналізу програмного комплексу що розроблюється, можна зробити висновок, що з альтернатив, що залишились після першого відбору двох варіантів виконання програмного комплексу оптимальним є другий варіант реалізації програмного продукту. У нього виявився найкращий показник техніко-економічного рівня якості $K_{\text{ТЕР}} = 3.61 \cdot 10^{-5}$.

Цей варіант реалізації програмного продукту має такі параметри:

- мова програмування – Python;
- середовище розробки JUPITER.

Даний варіант виконання програмного комплексу дає користувачу зручний інтерфейс, непоганий функціонал і швидкодію.

ВИСНОВКИ

В данній роботі було ознайомлення зі станом проблеми за обраною темою, збирання фактичних матеріалів, досліджень, покладена ідея програмного продукту.

У першому розділі описані теоретичні відомості телекомунікаційної сфери, визначення проблем, дослідження ситуації на ринку України і світу.

У другому розділі описані теоретичні аспекти моделювання процесів, їх представлення у формі часових рядів. Проведено огляд інформаційно-аналітичної системи. Також було обрано підприємство для дослідження, а саме – ПАТ «Укртелеком». Були переглянуті та проаналізовані існуючі методи і моделі прогнозування. Вибрано програмне забезпечення для первинної обробки даних, щоб потім на основі отриманих результатів побудувати математичну модель.

У третьому розділі було представлено програмний продукт, який може стати альтернативою для існуючих програмних забезпечень, а також може бути використаний на будь-якій платформі. Можна порівняти результати, отримані у другому розділі за допомогою ПЗ Eviews та моїм програмним продуктом.

У четвертому розділі описано результати функціонально-вартісного аналізу програмного продукту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Бідюк П.І., Коршевніук Л.О. Проектування комп'ютерних інформаційних систем підтримки прийняття рішень: навч. посіб. ННК «Інститут прикладного системного аналізу» Національний технічний університет України «Київський політехнічний інститут», Київ, 2010. 340 с.
2. Бідюк П.І. Часові ряди: моделювання та прогнозування, Київ: ЕКМО, 2004. 144 с.
3. Дрейпер Н., Смит Г. Финансы и статистика: в 2-х кн. Москва: 1986. Т.2: Прикладной регрессионный анализ. 366 с.
4. Бідюк П.І, Романенко В.Д., Тимошук О.Л. Аналіз часових рядів: навч. посіб., Київ: НТУУ «КПІ», 2013, 600 с.
5. Бідюк П. І. Система підтримки прийняття рішень для аналізу і прогнозування стану підприємства: навч посіб., Київ, 2013. с. 128–136.
6. Довгий С.Г., Бідюк П.І., Трофимчук О.М. Системи підтримки прийняття рішень на основі ймовірісно-статистичних методів., Київ: Азимут-Україна, 2014. 450 с.
7. Половцев О. В. Системний підхід до моделювання, прогнозування та управління фінансово-економічними процесами, О. В. Половцев, П. І. Бідюк, Л. О. Коршевніук, Донецьк, Східний видавничий дім, 2009, 286 с.
8. Бідюк П. І. Методи прогнозування, П. І. Бідюк, О. С. Меньяйленко, О. В. Половцев. – Луганськ : Альма Матер, 2008, 608 с.
9. Тихонов Е. Е. Методы прогнозирования в условиях рынка, Е. Е. Тихонов, Невинномысск, 2006, 221с.

ДОДАТОК А

Моделювання і прогнозування процесів у телекомунікаціях

Виконала: Абдал-Бакі А.М.

Науковий керівник: д.т.н. проф. Бідюк П.І.

Актуальність дослідження

Сектор телекомунікацій складається з компаній, які роблять можливим спілкування в глобальному масштабі, по телефону або через Інтернет, через ефір або кабелі, по дротах або по бездротовому зв'язку. Ці компанії створили інфраструктуру, яка дозволяє відправляти дані в вигляді слів, голосу, аудіо або відео в будь-яку точку світу. Найбільшими компаніями в цьому секторі є оператори телефонного (як дротового, так і бездротового) зв'язку, супутникові компанії, кабельні компанії і постачальники інтернет-послуг.

Не так давно телекомунікаційний сектор складався з клубу великих національних і регіональних операторів. З початку 2000-х років галузь була охоплена швидкою дерегуляцією і інноваціями. У багатьох країнах світу державні монополії в даний час приватизовані і стикаються з безліччю нових конкурентів. Традиційні ринки були перевернуті з ніг на голову, оскільки зростання мобільних послуг випереджає фіксовану лінію, і інтернет починає замінювати голос, як основний бізнес.

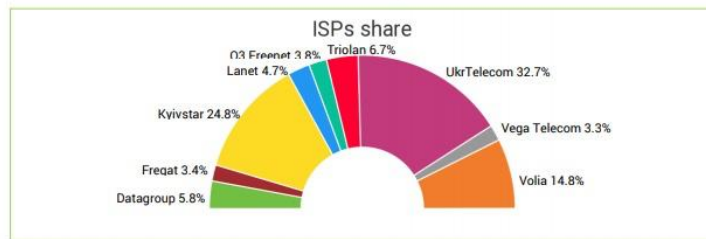
Огляд телекомунікаційного сектора

Аналітики передбачають, що інновації в продуктах і збільшення злиттів і поглинань будуть тільки сприяти подальшому зростанню і успіху телекомунікаційної галузі. Є багато можливостей для інвесторів, і збільшення інвесторів буде тільки сприяти подальшому розвитку сектора.

3% - середньорічний історичний темп зростання телекомунікаційного сектора.

Стабільність зростання сектора, навіть в періоди спаду, означає, що він вважається надійною захисною інвестицією, зберігаючи при цьому свою привабливість для зростаючих інвесторів. Навіть в нестабільні економічні часи стійкий попит на послуги передачі голосу і даних забезпечує стабільне джерело доходів для великих телекомунікаційних компаній.

Рейтинг українських інтернет-провайдерів



More than 70% of our tests were conducted on the networks of Kyivstar, UkrTelecom and Volia.

Французька компанія nPerf, яка спеціалізується на вимірюванні швидкості і якості доступу до мережі інтернет, опублікувала дані тестування мереж українських ШПД-провайдерів. Компанія має в своєму розпорядженні вимірювальні сервери на території України, за допомогою яких за 2018 рік було проведено 437 305 тестів підключення на дев'яти найбільших інтернет-мережах України.

Вибір підприємства для дослідження: ПАТ «Укртелеком»

Публічне акціонерне товариство «Укртелеком» — найбільший оператор фіксованого зв'язку в Україні. Компанія надає широкий спектр телекомунікаційних послуг в усіх регіонах країни. Уже багато років її справедливо вважають лідером серед провайдерів фіксованого інтернету.

Укртелеком має найбільше інтернет-покриття та найрозгалуженішу оптичну мережу.

У 2011 році компанія заснувала оператора «ТриМоб», який надає послуги мобільного зв'язку за технологією UMTS/WCDMA.

У 2015 році Укртелеком розпочав масштабну модернізацію мережі.

Укртелеком пропонує практично всі види сучасних телекомунікаційних послуг.

Побудова ARMA моделі на існуючих даних за допомогою Python

Позначення $AR(p, q)$ стосується моделі з p авторегресійними членами та q членами ковзного середнього. Ця модель містить моделі $AR(p)$ та

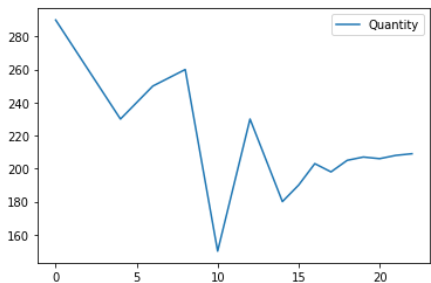
$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}.$$

Загальну модель $AR(p, q)$ було описано 1951 року в дисертації Пітера Уїттла, який використовував математичний аналіз (ряд Лорана та аналіз Фур'є) та статистичне висновування. Моделі $AR(p, q)$ було популяризовано книгою 1970 року Джорджа Бокса та Дженкінса, які виклали ітераційний метод (Бокса-Дженкінса) для їхнього вибирання та оцінювання. Цей метод був корисним для многочленів нижчих порядків (третього або нижчого ступеня).

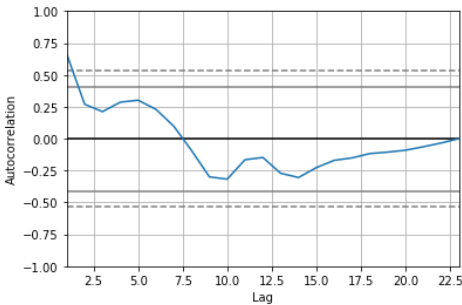
Ми маємо декілька видів даних:

- B2C (відключення/підключення фіксованого зв'язку та відключення/підключення інтернету; дані за останній рік)
- B2B (відключення/підключення фіксованого зв'язку та відключення/підключення інтернету; дані за останній рік)

Підключення фіксованого зв'язку B2C



Побудова тренду



Побудова автокореляції

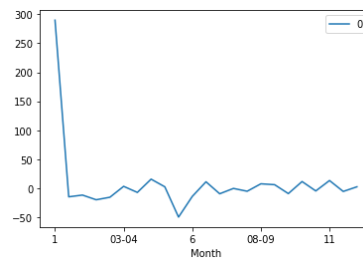
Підключення фіксованого зв'язку B2C

| SARIMAX Results | | | | | | |
|-------------------------|------------------|-------------------|---------|-------|--------|---------|
| ===== | | | | | | |
| Dep. Variable: | Quantity | No. Observations: | 23 | | | |
| Model: | ARIMA(5, 1, 0) | Log Likelihood | -90.824 | | | |
| Date: | Tue, 02 Jun 2020 | AIC | 193.647 | | | |
| Time: | 01:33:47 | BIC | 200.193 | | | |
| Sample: | 0 | HQIC | 195.189 | | | |
| | - 23 | | | | | |
| Covariance Type: | opg | | | | | |
| ===== | | | | | | |
| | coef | std err | z | P> z | [0.025 | 0.975] |
| ar.L1 | 0.6473 | 0.244 | 2.648 | 0.008 | 0.168 | 1.126 |
| ar.L2 | -1.1166 | 0.293 | -3.811 | 0.000 | -1.691 | -0.542 |
| ar.L3 | 0.6580 | 0.454 | 1.448 | 0.148 | -0.233 | 1.549 |
| ar.L4 | -0.5309 | 0.373 | -1.423 | 0.155 | -1.262 | 0.200 |
| ar.L5 | 0.2593 | 0.323 | 0.802 | 0.423 | -0.375 | 0.893 |
| sigma2 | 202.9423 | 61.523 | 3.299 | 0.001 | 82.359 | 323.525 |
| ===== | | | | | | |
| Ljung-Box (Q): | 4.31 | Jarque-Bera (JB): | 18.12 | | | |
| Prob(Q): | 1.00 | Prob(JB): | 0.00 | | | |
| Heteroskedasticity (H): | 0.51 | Skew: | -1.44 | | | |
| Prob(H) (two-sided): | 0.39 | Kurtosis: | 6.38 | | | |

Будуємо модель ARIMA (5,1,0). Встановлюємо значення відставання в 5 для авторегресії, використовуємо різницю порядку 1, щоб зробити часовий ряд нерухомим, і використовуємо ковзну середню модель 0.

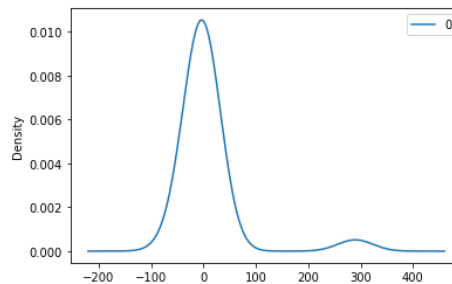
Підключення фіксованого зв'язку B2C

По-перше, ми отримуємо лінійний графік залишкових помилок, , що дозволяє припустити, що все ще може бути якась інформація про тренд, не зафіксована моделлю.



Підключення фіксованого зв'язку B2C

Далі ми отримуємо графік щільності залишкових значень помилок, припускаючи, що помилки є гауссовими, але не можуть бути зосереджені на нулі.



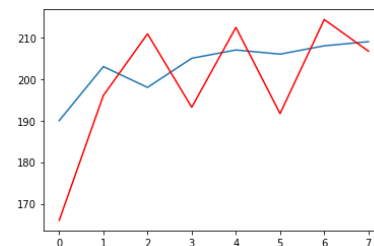
Підключення фіксованого зв'язку B2C

Відображається розподіл залишкових помилок. Результати показують, що дійсно існує упередженість у прогнозуванні (ненульове значення в залишках).

```
count    23.000000
mean      8.321421
std       63.007656
min       -50.035135
25%      -10.884624
50%       -4.876546
75%       6.684970
max       290.000000
```

Підключення фіксованого зв'язку B2C

- Ми можемо обчислити підсумковий середньоквадратичний показник помилок (MSE) для прогнозів, забезпечуючи точку порівняння для інших конфігурацій ARIMA.
- Створюється лінійний графік, який показує очікувані значення (синім кольором) в порівнянні з прогнозами змінного прогнозу (червоним кольором). Ми можемо бачити, що значення показують деяку тенденцію і знаходяться в правильному масштабі.



ДОДАТОК Б

```

from pandas import read_csv
from pandas import datetime
from matplotlib import pyplot
from pandas.plotting import autocorrelation_plot
from statsmodels.tsa.arima.model import ARIMA
from sklearn.metrics import mean_squared_error
import pandas as pd

def parser(x):
    return datetime.strptime('190'+x, '%Y-%m')

series = read_csv('B2CTELOFF.csv')
series.Month = series.Month.astype(str)
print(series.head())
series.plot()
pyplot.show()

series = series.set_index('Month')

autocorrelation_plot(series)
pyplot.show()

# fit model
model = ARIMA(series, order=(5,1,0))
model_fit = model.fit()
print(model_fit.summary())

```

```
# plot residual errors
residuals = pd.DataFrame(model_fit.resid)
residuals.plot()
pyplot.show()
```

```
residuals.plot(kind='kde')
pyplot.show()
print(residuals.describe())
```

```
X = series.values
size = int(len(X) * 0.66)
train, test = X[0:size], X[size:len(X)]
history = [x for x in train]
predictions = list()
for t in range(len(test)):
    model = ARIMA(history, order=(5,1,0))
    model_fit = model.fit()
    output = model_fit.forecast()
    yhat = output[0]
    predictions.append(yhat)
    obs = test[t]
    history.append(obs)
    print('predicted=%f, expected=%f' % (yhat, obs))

error = mean_squared_error(test, predictions)
print('Test MSE: %.3f' % error)
```

```
# plot  
pyplot.plot(test)  
pyplot.plot(predictions, color='red')  
pyplot.show()
```


ДОДАТОК В

```
from pandas import read_csv
from matplotlib import pyplot
from pandas import DataFrame
from pandas import concat
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.ar_model import AutoReg
from math import sqrt

series = read_csv('B2CTELOFF.csv', header=0, index_col=0)
print(series.head())
series.plot()
pyplot.show()

from pandas.plotting import lag_plot, autocorrelation_plot
lag_plot(series)
pyplot.show()

values = DataFrame(series.values)
dataframe = concat([values.shift(1), values], axis=1)
dataframe.columns = ['t-1', 't+1']
result = dataframe.corr()
print(result)

autocorrelation_plot(series)
pyplot.show()
```

```

from statsmodels.graphics.tsaplots import plot_acf
plot_acf(series, lags=8)
pyplot.show()

# create lagged dataset
values = DataFrame(series.values)
dataframe = concat([values.shift(1), values], axis=1)
dataframe.columns = ['t-1', 't+1']
# split into train and test sets
X = dataframe.values
train, test = X[1:len(X)-6], X[len(X)-6:]
train_X, train_y = train[:,0], train[:,1]
test_X, test_y = test[:,0], test[:,1]

# persistence model
def model_persistence(x):
    return x

# walk-forward validation
predictions = list()
for x in test_X:
    yhat = model_persistence(x)
    predictions.append(yhat)
test_score = mean_squared_error(test_y, predictions)
print('Test MSE: %.3f' % test_score)
# plot predictions vs expected
pyplot.plot(test_y)
pyplot.plot(predictions, color='red')

```

```
pyplot.show()
```

```
# split dataset
```

```
X = series.values
```

```
train, test = X[1:len(X)-6], X[len(X)-6:]
```

```
# train autoregression
```

```
window = 7
```

```
model = AutoReg(train, lags=7)
```

```
model_fit = model.fit()
```

```
coef = model_fit.params
```

```
# walk forward over time steps in test
```

```
history = train[len(train)-window:]
```

```
history = [history[i] for i in range(len(history))]
```

```
predictions = list()
```

```
for t in range(len(test)):
```

```
    length = len(history)
```

```
    lag = [history[i] for i in range(length-window,length)]
```

```
    yhat = coef[0]
```

```
    for d in range(window):
```

```
        yhat += coef[d+1] * lag[length-d-1]
```

```
    obs = test[t]
```

```
    predictions.append(yhat)
```

```
    history.append(obs)
```

```
    print('predicted=%f, expected=%f' % (yhat, obs))
```

```
rmse = sqrt(mean_squared_error(test, predictions))
```

```
print('Test RMSE: %.3f' % rmse)
```

```
# plot  
pyplot.plot(test)  
pyplot.plot(predictions, color='red')  
pyplot.show()
```